NEUROPSYCHOLOGICAL FUNCTIONING IN PEDIATRIC MULTIPLE SCLEROSIS: A LONGITUDINAL INVESTIGATION OF COGNITIVE PERFORMANCE AND THE CONTRIBUTION OF PRIMARY CLINICAL FACTORS

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

IN THE GRADUATE SCHOOL OF THE

TEXAS WOMAN'S UNIVERSITY

DEPARTMENT OF PSYCHOLOGY AND PHILOSOPHY

COLLEGE OF ARTS AND SCIENCES

BY

ANGELA M. CANAS, B.A., M.A.

DENTON, TEXAS

AUGUST 2014

DEDICATION

This accomplishment is dedicated to my late father, Gustavo A. Cañas Roldan. May you know that despite our short-lived years together, your nurturance and guidance have been ever-present in my life. It is my hope that the culmination of my labor and perseverance has validated your faith in me, and that your spirit will have peace in knowing the tenacious woman you created.

ACKNOWLEDGEMENTS

An accomplishment of this magnitude would not have been possible without the guidance, support, and encouragement of many individuals. I am foremost grateful to my parents, Elizabeth Espinoza and Jorge Chavarria, who have instilled in me the values of work ethic, perseverance, and scholarship. The sacrifices you made to foster my love of learning have played an integral role in the achievement of this project. I must also acknowledge my life partner, Jeff Lee, whose ability to make me laugh in times of distress is unparalleled. No other person knows quite as well as you the labor and sacrifice that I have put forth in this process. Thank you for allowing me to incorporate my love of learning in the bond that we share. I would also like to thank my siblings, Sylvia, Katherine, and Natally, and my niece, Monica, for their unconditional love and patience. Although this process has come with a myriad of challenges, I have always been able to count on your steadfast belief in my ability to succeed.

I would like to extend my deepest gratitude to Dr. Lana Harder for introducing me to the world of pediatric multiple sclerosis. Your genuine enthusiasm, generosity, and invaluable knowledge have been fundamental to my professional growth and enduring passion for the field of neuropsychology. I am tremendously fortunate to have had such an empathic and selfless mentor throughout this process. I am also indebted to Dr. Denise Maricle for always encouraging me to pursue training opportunities that would enrich my personal and professional development. Your unwavering academic guidance has been

fundamental in shaping the psychologist that I aspire to become. I am expressly grateful to Dr. Benjamin Greenberg, who established the Pediatric Demyelinating Diseases Clinic at Children's Medical Center of Dallas. Your expertise and thoughtful contributions have transformed my study into a true artifact of interdisciplinary collaboration. This accomplishment was also made possible through the dedication and supportive feedback of Drs. Daniel Miller and Wendi Johnson. Your reassurance and flexibility have provided me with a much needed sense of comfort.

The appreciation I have for the many friends and colleagues who have accompanied me along this journey cannot be overstated. I am particularly blessed to have my dear confidant and research counterpart, Cristina Sevadjian, in my life. Your genuine compassion and patience has impacted me in ways words cannot express. Thank you for helping me find the forest through the trees in times of need, for encouraging me to live in the moment, and for always providing me with unconditional positive regard. I am also thankful to Ashley Fournier-Goodnight for her candidness and camaraderie. No matter what the circumstance or question, you were always willing to help, and for that I am grateful. Lastly, I would like to thank Joanna and Phillip Morris for never failing to celebrate who I am and who I aspire to become.

ABSTRACT

ANGELA M. CANAS

NEUROPSYCHOLOGICAL FUNCTIONING IN PEDIATRIC MULTIPLE SCLEROSIS: A LONGITUDINAL INVESTIGATION OF COGNITIVE PERFORMANCE AND THE CONTRIBUTION OF PRIMARY CLINICAL FACTORS

AUGUST 2014

Multiple sclerosis (MS) is an autoimmune disease of the central nervous system that has been historically regarded as an adult medical condition. In recent years, however, the occurrence of MS among pediatric patients has been increasingly recognized, with an estimated one-third of patients exhibiting significant cognitive impairments. With that said, research devoted to understanding the nature, prevalence, and severity of MS symptomology in pediatric patients remains limited. To date, only two longitudinal studies exploring the evolution of cognitive functioning among pediatric cohorts have been published. Moreover, while adult studies have investigated the impact of neurologic impairment, disease duration, and relapse rate on cognitive performance, the relationship between primary clinical factors and cognitive dysfunction in pediatric patients remains elusive.

The present study assessed individual and group-level changes in cognitive functioning in a cohort of pediatric MS patients (n = 20) across a mean period of 14.65 months using a brief neuropsychological battery. The study employed the Reliable

Change Index (RCI) and Binomial Probability Distribution (BPD) methods in addition to a repeated measures multivariate analysis of variance (MANOVA) to accomplish this goal. The study also explored the relationship between cognitive performance and various clinical factors, including age at onset, relapse rate, disease duration, fatigue, and depression using a stepwise linear regression analysis. Examination of individual cognitive trajectories revealed more improvements than declines in performance, although 45% of participants met criteria for significant cognitive impairment. Declines in working memory, visual scanning and sustained attention, immediate verbal learning, and visual perception were most pronounced; whereas, motor-based processing speed and verbal fluency were largely preserved. Group-level analyses revealed significant declines in working memory, with noted improvements in visual scanning and motor speed. Several relevant clinical variables were noted to predict performance on Digit Span, SDMT, CVLT, TMT A, GPT, and Beery VMI. The directionality and interpretability of these predictive relationships, however, varied widely. Taken together, this study contributes to what is known about the impact of pediatric MS and associated clinical factors on cognitive functioning across time. Results support the use of routine neuropsychological screenings to monitor functioning and inform educational and treatment planning.

TABLE OF CONTENTS

Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

Multiple sclerosis, derived from the German term "multiplen sklerose," has been named and described in various ways over the last two centuries (Murray, 2005). Carswell, in one of the earliest illustrations of the disease, referred to it as a "peculiar disease state" (Carswell, 1838). Charcot, on the other hand, impressed by the distribution of sclerotic lesions, used the term *sclérose en plaque disseminée*, while the English preferred the term *insular sclerosis* (Murray, 2005; 2009). By 1899, the terms *multiple sclerosis*, *insular sclerosis*, *disseminated sclerosis, and cerebral sclerosis* were used interchangeably in textbooks. In 1921, the Association for Research in Nervous and Mental Diseases used both *multiple sclerosis* and *disseminated sclerosis* in the title of their publication. The term *multiple sclerosis*, however, became widely accepted in the 1950s, after being adopted by book authors and newly formed MS societies. By 1965, this term was embedded in virtually all English publications (Murray, 2005).

Multiple sclerosis (MS) is an autoimmune disease of the central nervous system (CNS) that affects approximately 1 in 800 individuals worldwide (Al-Futaisi, 2007; Ruggieri, Plasmati, & Simone, 2011). Pathologically, MS is characterized by a proliferation of sclerotic lesions resulting from episodes of inflammation and demyelination in subcortical white matter regions as well as more subtle tissue damage in diffuse areas of cortical grey matter (Fielding, Kilpatrick, Millist, & White, 2009).

Although the course of MS is largely heterogeneous, the disease is primarily characterized by unpredictable periods of acute symptom exacerbation (i.e., relapses) followed by periods of minimal disease activity (i.e., remission; Compston & Coles, 2008; Zanni, 2009). Clinical features often include optic nerve dysfunction (e.g., optic neuritis), sensory-motor disturbances, brainstem abnormalities, and bladder dysfunction (Banwell, Ghezzi, et al., 2007; Ghezzi & Banwell, 2011). Chronic symptoms are also prevalent and include cognitive impairment, psychological distress (e.g., depression, anxiety), and fatigue. These symptoms can be functionally impairing, often impacting educational performance, employment, and overall quality of life (Ghezzi & Banwell, 2011).

Although the classification of MS as an identifiable medical condition was largely shaped by the inquisitive landscape of the nineteenth century, early research focused primarily on the overt physical symptoms of the disease, which were believed to cause more severe impairment when compared to its cognitive and affective counterparts (Richardson, Robinson, & Robinson, 1997). While direct references to cognitive processes were rare, patients in the late 1800s were increasingly described as having an undifferentiated host of problems in the areas of fatigue, emotion, and cognition referred to as "mental symptoms" (Richardson et al., 1997). Charcot, for instance, noted a "marked enfeeblement of the memory" accompanied by "emotional faculties that are blunted in their totality" (Charcot, 1877, pp. 194-195). The lack of a more detailed and substantive commentary on cognitive functioning in such reports was largely attributed to

the fact that MS had been the object of clinical discussion for only a decade. Moreover, the perception that affective and intellectual symptoms were interdependent entities dominated the scientific realm of research at the time (Richardson et al., 1997).

The first series of thoughtful discussions and debates regarding the status of intellectual impairment in MS emerged in the 1920s (Richardson et al., 1997). During this period, researchers like Jelliffe and White (1919) began to stress the idea that mental symptoms in MS occurred more frequently than had been previously believed. Symptoms were more vividly described as "intellectual reductions in the form of hallucinatory states with mild confusion, difficulty in thinking, intermittent alterations in the capacity for attention and concentration, and lapses of memory" (Jellife & White, 1919, p. 517). An increase in the dissemination of similar reports reflected the burgeoning interest in the significance of MS symptomology, and more importantly, in how agreement with regard to appropriate methods for studying these symptoms might be reached (Richardson et al., 1997). Many researchers, in fact, began to move beyond the classic case study approach, which had dominated earlier investigations. Bohming, Cottrell, Wilson, and Ombredane played a critical role in shifting the emphasis of MS research away from qualitative case studies and toward the quantitative analysis of cohorts. Ross and Reitan, on the other hand, attempted to use more sophisticated tests to unravel the term "mental symptoms" and distinguish between intellectual and affective outcomes (Richardson et al., 1997).

A small, albeit significant, number of investigations on MS conducted during the late 1900s followed Ross and Reitan in stressing the importance of establishing the

relative independence between affective and cognitive symptomology (Richardson et al., 1997). While these investigations began to elucidate the impact of MS on cognitive functioning, it was the increasingly important role of formal neuropsychological assessment, the introduction and deployment of brain imaging techniques, and the growing interest in rehabilitative practices that propelled these studies to the forefront. An equally contributory factor was the emergence of the "cognitive revolution," an intellectual movement in both psychology and neuroscience that aimed to characterize the processes underlying intellectual behavior (Richardson et al., 1997).

Since these developments, cognitive impairment in adults with MS has become increasingly documented and quantified. Neuropsychological deficits are now widely recognized as primary and often disabling aspects of the disease process. Although a uniform pattern of cognitive impairment has not been established, various domains have been identified as common targets. Specifically, it appears that overall intelligence, primary language functions, and attention span remain relatively intact; however, speed of information processing, complex attention, conceptual reasoning, and various facets of memory are often impaired (Amato, Zipoli, & Portaccio, 2008). The greatest deficits are usually found in processing speed, learning, and memory, faculties that are largely affected by myelin depletion (Fields, 2008; Nagy, Westerberg, & Klingberg, 2004).

Improvements in psychometric methodologies coupled with scientific inquisition transformed the "mental symptoms" of the 1800s into the more narrowly defined psychological consequences of the twenty-first century. The cognitive deficits associated

with MS have certainly received increasing attention, making MS both a neurologically complex and psychologically involved condition. Unfortunately, the manifestation of these deficits has been predominantly acknowledged in and reserved for adult patients, despite growing evidence of a relatively comparable pediatric disease course. This void has been the culmination of various factors including the rarity of pediatric MS, the preponderance of attention on adult patients, the almost exclusive use of single case studies and small clinical samples, and the lack of variability in measurement techniques (Julian, Trojano, Amato, & Krupp, 2011).

The literature on the neuropsychological implications of pediatric MS continues to reflect the impact of these factors. While researchers have deviated from utilizing case studies to illustrate deficits on global measures of intelligence, comprehensive, cohort-based studies have only begun to emerge within the last decade. The systematic investigation conducted by Banwell and Anderson in 2005, in fact, was the first of its kind, and only a limited number of cross-sectional studies (i.e., Amato, Gorretti, et al., 2008; MacAllister et al., 2005; Portaccio et al., 2009) have followed. Collectively, these studies have demonstrated great overlap in cognitive dysfunction between adult and pediatric MS patients. Specifically, involvement of memory functions, complex attention, processing speed, and various aspects of executive functions appear to be implicated across all ages. Despite these findings, several questions regarding the long-term progression of cognitive dysfunction in children and adolescents have yet to be answered. The paucity of longitudinal research is exemplified by the fact that only two studies

exploring the neuropsychological outcomes of pediatric MS over time have been published. Moreover, while adult studies have alluded to the impact of neurologic impairment (i.e., EDSS), disease duration, and relapse rate on cognitive performance, the relationship between primary clinical factors and cognitive dysfunction in pediatric patients remains elusive (Julian et al., 2011).

Purpose, Rationale, and Significance of the Current Study

The primary purpose of the present study was to assess cognitive functioning in a cohort of children and adolescents with MS over time. Although two relatively recent longitudinal studies (Amato et al., 2010; MacAllister, Christodoulou, et al., 2007) have roused the interests of researchers in the field, the literature with respect to the evolution of cognitive deficits in MS patients has remained rather limited. Existent longitudinal studies have primarily utilized a frequency-based approach emphasizing the number of patients performing in the impaired range or the number of failed tests at follow-up relative to baseline. Of note, time between baseline and follow-up has been relatively similar (21.58 to 24 months) across studies. Results from these studies suggest that pediatric MS patients exhibit significant cognitive declines in verbal learning and memory, complex attention, and verbal fluency within a relatively short period of time (Amato et al., 2010; MacAllister, Christodoulou, et al., 2007). However, to better understand how MS impacts the cognitive functioning and overall prognosis of pediatric MS patients, studies should emphasize both individual and group-based changes across varying time points. The present study responded to this need by examining changes in

performance across a 12 to 18 month period at both the individual and group levels. Of note, minimizing the period between baseline and follow-up has the potential for decreasing the extent to which findings are attributable to extraneous variables, such as treatment effects, rather than true changes in functioning. Moreover, if findings suggest that pediatric patients with MS experience some degree of decline within this relatively short period of time, the need for regular (i.e., annual) screenings may be supported. Such findings are expected to have important implications regarding standard of care for this rare disease group. These findings, for instance, could be used to dissuade insurance providers from extending suggested timelines for reevaluation.

Given the progressive nature of this disease, determining whether cognitive performance remains stable or changes over time is valuable for various reasons. Foremost, it is currently difficult to draw broad conclusions from the limited number of published longitudinal studies. This study has the potential for enhancing the literature base from which to extrapolate patterns of performance and information regarding the relationship between time and cognition. Second, the onset of pediatric MS can occur during significant developmental stages that impact academic, social, and daily functioning; thus, identifying appropriate interventions is critical to the care of children and adolescents with MS (Julian et al., 2011). Identifying a pattern or rate of progression associated with cognitive deficits in specific domains could inform educational and treatment planning. Lastly, the effects of demyelination of the white matter pathways in the developing brain have been postulated to differ from those of the mature adult brain

(Ross, Schwebel, Rinker, Ness, & Ackerson, 2010). While comparisons between children and adults have been presented in the literature, this study could help to further delineate the cognitive domains most vulnerable to disruption during active neurodevelopment.

This study was also conducted in order to elucidate the relationship between neurocognitive functioning and various clinical factors, including age at onset, relapse rate, disease duration, fatigue, and depression. Particularly, the study assessed whether, and to what extent, changes in neurocognitive performance correlated with the aforementioned factors. While poor cognitive performance has been associated with higher levels of neurologic dysfunction as a consequence of higher relapse rates and longer disease duration (MacAllister et al., 2005), the relationship between cognitive dysfunction and these clinical features in pediatric MS has remained largely elusive (Julian et al., 2011). Moreover, depression and fatigue have been noted as two of the most commonly experienced and debilitating symptom in adults with MS, significantly impacting daily functioning, productivity, and overall quality of life (Feinstein, 2011; Julian et al., 2011; MacAllister et al., 2009). However, the contribution of these factors to the cognitive deficits experienced by pediatric MS patients has hardly been explored. Addressing this void is particularly important given the potential differential impact of these factors on treatment outcomes, educational performance, and overall quality of life in pediatric patients with MS.

Research Questions and Hypotheses

Based upon a review of the literature, which establishes a rationale for this study, the following research questions were addressed:

1. Do pediatric MS patients experience changes in cognitive functioning, as measured by performance on specific standardized tests, over a mean period of 14.65 months?
2. Is there a significant predictive relationship between relevant clinical factors (i.e., age at onset, disease duration, relapse rate, fatigue, depression) and change in neurocognitive performance over time?

With regard to the former research question, it was hypothesized that the sample of pediatric MS patients in the study would demonstrate declines in verbal fluency, complex attention (i.e., sustained, shifting), and verbal learning and memory. This hypothesis was based on the fact that declines in these cognitive domains have been consistently evidenced across previous longitudinal studies. With regard to the latter research question, it was hypothesized that depression and fatigue would be negatively correlated with cognitive performance over time; whereas, a positive correlation between age at onset and cognitive functioning would be evidenced. Depression was selected as a variable of interest based on prior research, which has demonstrated an inverse relationship between depressive symptomology and planning ability, working memory capacity, and attentional processes (Julian & Arnett, 2009; Julian et al., 2011). Fatigue was chosen as a variable of interest given its complex pathophysiology and impact on the

following systems: CNS, immune, neuroendocrine, and neurotransmitter (MacAllister & Krupp, 2005). Younger age at onset was selected as a variable of interest based on developmental considerations. Specifically, younger age has been associated with the presence of immature pathways and developing networks (Fields, 2008; Julian et al., 2011). The disruption caused by MS to these developing networks and pathways can result in significant cognitive dysfunction and limitations in skill acquisition (Fields, 2008; Julian et al., 2011; MacAllister et al., 2012).

CHAPTER II

LITERATURE REVIEW

Since the differentiation and framing of MS as a separate identifiable entity in the nineteenth century (Murray, 2009), the medical field has witnessed a surge of interest and sensitivity to the issues surrounding the diagnosis of pediatric MS (Chabas, Krupp, & Tardieu, 2011). While such a concentrated interest in pediatric MS is relatively new, increased clinical awareness has translated into growing research efforts and publications among the international community. The continuous dissemination of information on pediatric MS has elucidated its debilitating effects and the importance of early identification and intervention (Chabas et al., 2011). With that said, much remains to be discovered with regard to the neuropsychological manifestation of MS over time (Renoux & Waubant, 2011).

This chapter provides a global overview of MS, with emphasis on its epidemiology and etiology. Moreover, a discussion on the clinical and neuroimaging features of MS will serve as a basis for discussing its diagnostic complexity. Additionally, this chapter will review the crucial role of myelin in cognitive development and the knowledge that has been gained from cross-sectional and longitudinal studies as it relates to the neuropsychological sequelae associated with pediatric MS. Relevant research related to the utilization of brief batteries in the assessment of pediatric MS will also be summarized. Finally, various measures designed to assess those processes with

purported susceptibility to MS, in addition to confounding factors that can affect the performance of MS patients on standardized measures, will be reviewed.

Overview of Multiple Sclerosis

Epidemiology

Incidence and prevalence. While the worldwide prevalence of MS is largely unknown, it is estimated to affect 1 in 800 individuals, or .12% of the general population, with noted regional and ethnic variability (Al-Futaisi, 2007; Ruggieri et al., 2011). In high-risk areas such as Canada and the United Kingdom (UK), for instance, the prevalence has been noted to exceed 200 per 100,000, although rates as low as 5 per 100,000 have been reported in regions such as the Middle East and the Caribbean (Marrie, Yu, Blanchard, Leung, & Elliott, 2010; Ruggieri et al., 2011). Despite the recent proliferation of research, the percentage of pediatric patients in the total MS population remains unclear (Banwell, Ghezzi, Bar-Or, Mikaeloff, & Tardieu, 2007). While population and registry-based studies have attempted to establish accurate rates of incidence and prevalence, data culled from such efforts have been plagued by various factors (Ruggieri et al., 2011). The absence of uniform and definitive criteria by which to establish consistent definitions of pediatric MS, for instance, have resulted in wide-ranging estimates (Al-Futaisi, 2007; Ruggieri et al., 2011). Moreover, given the rarity of this condition, well-developed studies comprised of clinical data gathered from an array of settings have been largely nonexistent, decreasing the likelihood of obtaining accurate rates (Ruggieri et al., 2011).

Within the context of such pitfalls, several retrospective studies (Boiko, Vorobeychik, Paty, Devonshire, & Sadovnick, 2002; Ferreira, Machado, Dantas, Moreira, & Souza, 2008) have estimated the overall prevalence of MS, with onset prior to the age of 18 years, to range from 1.6 to 10.5%. Incidence rates are equally variable with estimates as high as 1.3 per 100,000 in Japan and as low as .3 per 100,000 in Germany (Pohl, Hennemuth, von Kries, & Hanefeld, 2007; Torisu et al., 2010). Within the United States, the overall prevalence rate has been estimated at .15%, with an incidence rate of 1 in 700; consequently, there are approximately 400,000 adults with the disease (Ascherio & Munger, 2010). Between 2.7 and 5% of this estimate represents individuals with onset prior to the age of 18 years, while those with onset prior to the age of 10 represent about .2 to .7% of the total number of estimated cases (Boiko, Vorobeychik, et al., 2002; Chitnis, Glanz, Jaffin, & Healy, 2009).

Gender. In the adult population, females are more often affected by MS than males, with the female-to-male ratio ranging from 1.7:1 to 3.2:1 (Polliack, Barak, & Achiron, 2000; Simone et al., 2002). While recent epidemiological studies have shown an overall widening gender gap, a more even distribution has been noted in adult patients with late-onset MS (Kis, Rumberg, & Berlit, 2008). In the pediatric population, the female-to-male ratio varies depending on the age at disease onset (Ruggieri et al., 2011). For those with onset prior to 6 years of age, a predominance of affected males is evident, with a ratio of 0.8:1 (Chitnis et al., 2009; Simone et al., 2002). The discrepancy is particularly prominent in rare instances of onset prior to 24 months of age (0:0.6; Cole,

Auchterlonie, & Best, 1995). Conversely, the ratio appears to favor females between the ages of 6 and 10 years (1.6:1) and those in pre-adolescence (2.1:1; Chitnis et al., 2009; Simone et al., 2002). An even higher female preponderance has been recorded for patients with onset during adolescence (i.e., 3.35:1 at age 13; 7.67:1 at age 14; Boiko, Vorobeychik, et al., 2002). Overall, females are at higher risk for developing MS during adolescence through young adulthood, while males are at higher risk during the younger pediatric and later adult stages (Ruggieri et al., 2011).

Age at onset. The overall age distribution of disease onset is bell shaped. That is, 10% of individuals develop their first symptoms prior to the age of 20 years, 70% between the ages of 20 and 40 years, and 20% after the age of 40 years (Ruggieri et al., 2011). Patients with pediatric-onset MS have been traditionally divided into four categories: extremely early onset MS (i.e., prior to 24 months of age), MS in preschool children (i.e., under 6 years of age), prepubertal MS (i.e., between the ages of 6 and 12 years of age), and postpubertal, juvenile, or adolescent MS (i.e., between 13 and 18 years of age; Ruggieri et al., 2011). While only a few cases of extremely early-onset MS have been reported, the disease accounts for 0.8 to 14%, 0.5 to 30%, and 40 to 80% of preschool, preadolescent, and adolescent cases, respectively (Boiko, Vorobeychik, et al., 2002; Simone et al., 2002; Mikaeloff, Caridade, Assi, Suissa, & Tardieu, 2006).

Race and ethnicity. Race and ethnicity appear to modify the incidence and characteristics of all individuals with MS, although the racial and ethnic distribution of the disease in the pediatric population appears to differ from that observed in adults

(Ruggieri et al., 2011; Yeh et al., 2009). MS in adults is most common in Caucasians, while African Americans and Asians are less often affected. Studies published across North American regions, however, have denoted marked racial and ethnic variability in pediatric samples (Yeh et al., 2009). A recent retrospective study (Chitnis et al., 2009) derived from an MS Center in Boston, for instance, indicated a higher proportion of African Americans in the pediatric-onset group when compared to the adult-onset group (7.4% versus 4.3%). This trend has also been demonstrated in multiple Canadian cohort studies (Banwell et al., 2009). Moreover, reports from the Pediatric MS Center in Long Island, New York, have indicated a significantly higher proportion of pediatric individuals of African American and Hispanic ethnicity (Krupp, McLinskey, & Troell, 2008).

Etiology

Pathogenesis. MS is a chronic autoimmune disease involving the inflammation, demyelination, and neurodegeneration of the central nervous system (CNS; Chitnis & Bar-Or, 2011). While prevailing theory emphasizes the importance of oligodendrocytic atrophy and subsequent myelin depletion, MS has a complex etiology founded on both genetic and environmental factors (Waubant & Chabas, 2009; Yeh et al., 2009). The disease process is generally characterized by the inability of the body's cells to recognize myelin as a familiar entity, thus leading to inflammation, prolonged immunological activation (i.e., infiltration of T-cells, B-cells, and lipid-lade microphages), and blood-brain barrier leakage (Compston & Coles, 2002; MacAllister et al., 2012).

Progressive thinning of the axonal sheath accompanied by neuronal damage results in subsequent gliosis, a proliferation of astrocytes that leads to the formation of glial scars, also referred to a sclerotic plaques or lesions (Compston & Coles, 2008). Given the detrimental effects of MS on the myelinated axons of nerve cells in the subcortical white matter of the brain, such perivascular sclerotic lesions have traditionally served as the hallmark of MS pathology (Chitnis & Bar-Or, 2011).

Lesion distribution has been primarily evidenced across the white matter of the cerebrum, brain stem, and cerebellum (Wingerchuck & Weinshenker, 2000), although specific regions appear to be more susceptible than others. These regions include, but are not limited to, the periventricular white matter, corpus callosum, corona radiata, internal capsule, centrum semiovale, and the visual pathways (Wingerchuck & Weinshenker, 2000). The promulgation of lesions, however, does not appear to be isolated to white matter structures. All myelinated structures, including the gray matter of the cerebral cortex and basal ganglia, have been increasingly implicated in MS pathology (Geurts et al., 2007; Vercellino et al., 2005). Gray matter lesions may, in fact, contribute to the wide-ranging spectrum of clinical presentation and symptomology observed in MS patients (Vercellino et al., 2005). Although the precise mechanisms underlying MS are not fully understood, sclerotic lesions appear to impact nerve transmission by slowing or blocking electrochemical conduction via action potential damage (Arrondo et al., 2009). Such abnormalities in nerve conduction result in the clinical manifestation of various physical, cognitive, and emotional symptoms (Arrondo et al., 2009).

Unfortunately, knowledge related to the biological underpinnings leading to the development of childhood MS is limited (Chitnis & Bar-Or, 2011). The few available cases with detailed pathology indicate "a dense accumulation of lymphocytes and macrophages in a prominent perivascular distribution, with rare B cells" (Chitnis & Bar-Or, 2011, p. 158). Additional information with regard to the pathophysiology of childhood MS has also emerged from recent studies (Banwell et al., 2008; O'Connor et al., 2010; O'Connor et al., 2007) focusing on irregularities in T-cell proliferation and cerebral spinal fluid elevations. Among a mixed-clinical sample of children with various CNS autoimmune diseases (e.g., MS, Type I diabetes), for instance, T-cell proliferation was directed against a wide range of non-invasive, self-antigens relative to controls. These children also exhibited abnormal T-cell responses against various cow-milk proteins (Banwell et al., 2008). Similar safeguards against myelin oligodendrocyte glycoprotein (MOG; O'Connor et al., 2007) and myelin basic protein (MBP; O'Connor et al., 2010) have been observed among cohorts of children with initial demyelinating events, although antibodies against the latter appear to be associated with an encephalopathic type of onset (O'Connor et al., 2010). Despite these findings, no systematic studies evaluating and comparing the physiology of pediatric MS against adult MS have been conducted (Chitnis & Bar-Or, 2011; Yeh et al., 2009).

Genetic underpinnings. Documented familial recurrence rates for MS range from 15 to 20%, although the risk seems to decrease from 3% in first-degree relatives to 1% in those of second and third-degree kinship (Compston & Coles, 2002, 2008).

Population-based studies (Mumford et al., 1994; Willer, Dyment, Risch, Sadovnick, & Ebers, 2003) conducted in Canada and the UK have demonstrated higher recurrence rates in monozygotic pairs, while investigations conducted in France and Italy (French Research Group on Multiple Sclerosis, 1992; Ristori et al., 2006) have documented equivalent rates regardless of zygosity. The risk for half-siblings and offspring of single-affected parents is lower than that of full siblings and children of conjugal parents (Compston & Coles, 2002, 2008). Conversely, children raised by an adoptive family with an affected member appear to exhibit the same risk as does the general population. Together, these studies suggest that familial clustering and individual susceptibility to MS follow a polygenic pattern (Comspton & Coles, 2002, 2008).

Investigative results from various adult population studies suggest an association between MS and the class II Major Histocompatability Complex alleles DR15 and DQ6 with the following corresponding genotypes: DRB1*1501, DRB5*0101, DQA1*0102, and DQB2*0602 (Boiko, Gusev, et al., 2002; Compston & Coles, 2002). While this association is strongest in individuals of northern European descent, it has been noted in various populations with the exception of Mediterranean groups (e.g., Sardinians), who appear to have a distinct DR3 and DR4 association with the following genotypes: DRB1*0405, DQA1*0301, and DQB1*0302 (Marrosu et al., 2004). While the proliferation of additional studies with promising findings has been largely absent (Compston & Coles, 2002, 2008), markers HLA-C5 and HLA-DRB1*11 have been

recently added to the list of susceptibility loci (Dean et al., 2008; Ramagopalan et al., 2007; Yeo et al., 2007).

Environmental risk factors. Various environmental factors have been shown to contribute to MS susceptibility (Waubant & Chabas, 2009; Yeh et al., 2009). The most relevant and widely studied factors include an individual's latitude of residence, levels of vitamin D, exposure to particular infectious agents, and smoking practices (Ruggieri et al., 2011; Yeh et al., 2009). With regard to residency, various studies clearly indicate that adult MS is a geographically-related disease with a latitudinal gradient and a critical age at migration of 15 years. Specifically, increased prevalence rates have been noted among populations living in regions that are geographically located further away from the equator (Ruggieri et al., 2011; Yeh et al., 2009). Data on the geographic distribution of MS in the pediatric population, however, is not as definitive, and whether it is comparable to the pattern of adult MS has not been established (Ruggieri et al., 2011).

The linkage between latitude and increased risk of MS has also led to the hypothesis of an inverse relationship between sunlight exposure and MS (Van der Mei et al., 2003). Specifically, Vitamin D, which is metabolized in the skin via ultraviolet irradiation, has been shown to have modulating effects, in such that its depletion has been linked to MS susceptibility in the adult population (Yeh et al., 2009). While this relationship has not been confirmed in children, sub-normal levels of 25-hydroxy vitamin D_3 have been noted in a vast majority of pediatric MS patients across clinics (Waubant, Chabas, Strober, et al., 2009).

Determining the relevance of viral exposure has proven difficult in adult MS cases, given the lapse in time between initial exposure and disease onset (Yeh et al., 2009). Inevitably, most individuals will have encountered a variety of infectious agents by the time they reach adulthood, although only a small fraction of the population will have developed MS, making it difficult to ascertain the relationships between the two variables (Yeh et al., 2009). Nevertheless, seroepidemiologic and pathologic studies have produced robust evidence suggesting that previous infections with members of the Herpesviridae family may be associated with adult-onset MS (Rodriguez-Violante, Ordonez, Bermudez, Sotelo, & Corona, 2009). Such viruses include the herpes simplex virus, vericella zoster virus, human herpesvirus-6, Epstein-Barr virus, and cytomegalovirus (Rodriguez-Violante et al., 2009). The most widely studied member of this family and its impact on the development of MS has been the Epstein-Barr virus (EBV; Ruggieri et al., 2011). Theories related to the pathogenic role of EBV in MS include antigen mimicry, cytotoxic T-cell dysfunction, and immortalization of B-cell clones (Ascherio & Munger, 2010; Pohl, 2009). While these theories lack unequivocal supportive data, some studies have demonstrated elevated titers of EBV antibodies in adults with MS years prior to the development of neurological symptoms (Jaquiery et al., 2010; Lunemann et al., 2008).

An interest in the contribution of viruses to the development of MS in the pediatric population has also been demonstrated in the literature. The pediatric population, in fact, provides a unique opportunity to study the association between virus

exposure and MS, given the close temporal relationship between time of infection and MS onset (MacAllister et al., 2012). Moreover, children are more likely to have limited exposure to a wide range of viruses in comparison to adults (Yeh et al., 2009). Banwell, Krupp, and colleagues (2007) conducted a retrospective study on 136 children with MS in North and South America and Europe to ascertain the relationship between viral infection and MS development. While differences between the MS group and their age-matched controls were not observed with regard to seropositivity to a multitude of viruses (e.g., cytomegalovirus virus, varicella zoster virus), EBV seropositivity was associated with a significant increase in likelihood of developing MS (Banwell, Krupp, et al., 2007). In a similar study, Pohl and colleagues (2010) analyzed the amount of cerebral spinal fluid antibody production against EBV, varicella zoster virus, and herpes simplex virus, in addition to measles and rubella, in a sample comprised of 43 pediatric-onset MS patients, 50 adult-onset patients, and controls. While seropositivity for EBV was more prevalent among the patients than the controls, frequency of antibody detection for all viruses was comparable (Pohl et al., 2010). Thus, the implication of EBV and other viruses in pediatric-onset MS remains tentative (Ruggieri et al., 2011).

In addition to viral exposure, concerns with regard to the utilization of vaccinations, namely the hepatitis B vaccine, and the subsequent development of MS have been raised (Yeh et al., 2009). To this effect, two case-control studies conducted in France (Touze, Gout, Verdier-Taillefer, Lyon-Caen, & Alperovitch, 2000; Touze et al., 2002) reported a 40 to 70% increased risk of developing MS within two months

post-immunization, while a study (Sturkenboom et al., 1999) conducted in the UK estimated a 60% increased risk within 12 months following vaccination. However, five additional studies (Ascherio et al., 2001; DeStefano et al., 2003; DeStefano, Weintraub, & Chen, 2005; Sadovnick & Scheifele, 2000; Zipp, Weil, & Einhaupl, 1999) were unable to reproduce such findings. Given the variability of research findings, the relationship between hepatitis B vaccination and adult MS remains highly controversial (Ruggieri et al., 2011).

To date, only a small cluster of studies (Mikaeloff, Caridade, Rossier, Suissa, & Tardieu, 2007; Mikaeloff, Caridade, Suissa, & Tardieu, 2009) conducted in France by the same group of authors have examined the role of the hepatitis B vaccine in pediatric MS. Across their studies, the authors failed to establish an association between hepatitis B vaccination and the development of MS over the course of 3 years post-vaccination. Moreover, the authors found no increase in the relapse rate of patients who had subsequently been vaccinated against hepatitis B. The authors reported, however, a trend for the hepatitis B vaccine to increase the risk of MS in the long-term (Mikaeloff, Caridade, Rossier et al., 2007; Mikaeloff et al., 2009).

In the adult population, active smokers have been noted to have a 40 to 80% higher risk of developing MS when compared to non-smokers (Hedstrom, Baarnhielm, Olsson, & Alfreddson, 2009; Riise, Norvtvedt, & Ascherio, 2003). The risks associated with cumulative doses of smoking include increased blood-brain barrier disruption, greater levels of atrophy, and higher lesion volumes (Zivadinov et al., 2009) with

subsequent adverse effects related to disease progression (Healy et al., 2009). These premises have served as the foundation for a case-controlled study investigating the linkage between exposure to passive smoking and childhood-onset MS (Mikaeloff, Caridade, Tardieu, Suissa, & KIDSEP study group, 2007). Results from this study suggest that children who have been exposed to parental smoking have a significantly higher risk of developing MS; this risk was noted to increase with prolonged exposure of 10 years or more. It is important to note, however, that a familial history of MS may have played a contributing role in these results (Mikaeloff, Caridade, Tardieu, et al., 2007). While the underlying mechanisms of active and passive smoking remain unclear, it is postulated that the direct exposure to cyanide and thiocyanate, which have been demonstrated to cause myelin depletion in the CNS of animal models (Bass, 1968) and obstruction of adenoid-mediated interferon gamma production in children (Avanzini et al., 2006), are at play.

Clinical Presentation and Diagnosis

Course. Although the course of MS is largely heterogeneous, four phenotypic subtypes have been identified: relapsing-remitting (RRMS), primary progressive (PPMS), secondary progressive (SPMS), and progressive-relapsing (PRMS; Zanni, 2009). RRMS is characterized by unpredictable periods of acute symptom exacerbation followed by periods of remission with minimal to no symptoms or signs of disease activity (Compston & Coles, 2008; Zanni, 2009). RRMS is the most common subtype, accounting for 80 to 85% of all MS initial diagnoses. Approximately 65% of individuals with an initial RRMS

diagnosis will gradually develop SPMS. This subtype is characterized by a similar waxing and waning pattern, although fewer relapses have been noted. Nevertheless, symptoms are usually more pronounced and higher degrees of impairment are evidenced (Zanni, 2009). PPMS represents 10 to 15% of cases and is characterized by an absence in remission following the initial episode of symptoms in addition to a steady worsening of symptoms across time (Waubant & Chabas, 2009; Zanni, 2009). PRMS is the rarest subtype, affecting 5% of individuals who experience progressive decline accompanied by occasional acute attacks (Zanni, 2009).

Although the overall course of MS is similar in children and adults, various differences have been noted. For instance, children almost exclusively present with a relapsing remitting course (i.e., 85.7 to 100% of cases); whereas, approximately 15% of adult-onset cases are initially diagnosed as PPMS (Boiko, Vorobeychik, et al., 2002; Waubant & Chabas, 2009). Within this context, the annual relapse rate is higher in children than adults, with estimates ranging from .38 to .87 (Renoux & Waubant, 2011). Moreover, a slightly shorter mean time between the first and second episode has been observed in pediatric patients (1.6 years) when compared to adult cohorts (1.7 to 2.0 years; Renoux & Waubant, 2011). Conversely, transition to SPMS seems to favor individuals with pediatric-onset MS, in that the course appears to be more gradual, taking a median of 23 to 28 years in comparison to 10 to 18 years in patients with adult-onset MS (Simone et al., 2002; Renoux et al., 2007). Despite having a more gradual course,

patients with pediatric-onset MS experience higher degrees of disability at a younger age (Renoux & Waubant, 2011).

Clinical features. Pediatric-onset MS is accompanied by a myriad of clinical features. Monofocal presentations often include optic neuritis, sensory-motor disturbances, brainstem abnormalities, and bladder dysfunction (Banwell, Ghezzi, et al., 2007; Ghezzi & Banwell, 2011). Approximately 5 to 65% of children experience sensory-motor disturbances, 3 to 20% have cerebellar symptoms, 13 to 40% present with brainstem involvement, and 12 to 37% are diagnosed with optic neuritis at onset; urinary symptoms, on the other hand, appear to be less common, affecting 1 to 9% of children (Ghezzi, 2004; Ghezzi & Banwell, 2011). Seizures occur in approximately 5% of pediatric cases, although they are almost exclusively implicated in children with onset prior to the age of 10 years (Ruggieri, Iannetti, Polizzi, Pavone & Grimaldi, 2004). Ataxia, encephalopathy, and fever also appear to affect children within this age group (Callen, Shroff, Branson, Lotze et al., 2009). Polyfocal presentations, however, are more common in pediatric MS relative to adult MS (Yeh et al., 2009), with such presentations representing at least half of all pediatric-onset cases (MacAllister et al., 2012). Polyfocal features are most notable in children with onset prior to the age of 6 years (Ruggieri et al., 2004), while monofocal features are most common in adolescents, particularly those from European countries, as compared to those from North and South American regions (Banwell, Krupp, et al., 2007). Chronic symptoms are also prevalent and include cognitive impairment and fatigue (Ghezzi & Banwell, 2011).

Neuroimaging features. In many respects, MRI visualization in children and adults is comparable, with a notable presence of periventricular lesions (Callen, Shroff, Branson, Lotze et al., 2009). Nevertheless, there are various distinct features that have been evidenced in the neuroimaging of pediatric cases. Children with MS, for instance, appear to have a higher disease burden at onset with a greater frequency of infratentorial lesions of the brainstem, pons, and cerebellum (Ghassemi et al., 2008; Waubant, Chabas, Okuda, et al., 2009). Differences in the location of pathology are likely associated with ongoing neurodevelopmental changes, such as those related to the caudal-to-rostral temporal gradient of myelination, although this hypothesis remains to be confirmed (Soares, Chabas, & Wintermark, 2011). A propensity for large demyelinating lesions with marked perilesional edema in addition to tumefactive plaques has also been reported in a number of children (McAdam, Blaser, & Banwell, 2002). Lastly, thoracic spinal nerve 2 (T2) lesion volumes in pediatric patients are often higher than those seen in adults who have comparable disease durations (Yeh et al., 2009). Within pediatric cohorts, differences in lesion characteristics have been noted to vary by age. Specifically, children with a pre-pubertal onset tend to have lesions that are amorphous, diffuse, and bilateral (Callen, Shroff, Branson, Lotze et al., 2009). In addition, these children are more likely to have fewer enhancing lesions and deeper gray matter involvement (Chabas et al., 2008). Lesions in younger cohorts also appear to be less consistent, with many children exhibiting a reduction in lesion activity on subsequent MRI scans (Callen, Shroff, Branson, Lotze et al., 2009; Chabas et al., 2008).

Diagnostic criteria. MRI diagnostic criteria for adult-onset MS are well-established and widely used (MacAllister et al., 2012). Historically, three or more of the following criteria have been required in order to demonstrate dissemination in space: a minimum of one gadolinium-enhancing lesion or nine T2 hyperintense lesions in the absence of a gandolinium-enhancing lesion; a minimum of one infratentorial lesion; a minimum of one juxtaposed lesion; a minimum of three periventricular lesions (McDonald et al., 2001; Polman et al., 2005). Furthermore, to demonstrate dissemination in space, the following criteria have been implemented: detection of gandolinium enhancement at least three months after onset of the initial event or detection of a new T2 lesion at any point beyond 30 days from onset of the initial event (McDonald et al., 2001; Polman et al., 2005). These criteria, however, have recently been revised to reflect a simplified process that allows for earlier diagnosis (Polman et al., 2011). With regard to dissemination in space, it is now required that a minimum of one or more T2 lesions be demonstrated in at least two of the following areas of the CNS: periventricular, juxtacortical, infratentorial, and spinal cord. With regard to dissemination in time, while an individual must still experience a minimum of two events accompanied by positive MRI findings, it is no longer necessary that the detection of lesions be separated by a period of 30 or more days (Polman et al., 2011).

Despite noted efforts, a set of well-developed and definitive criteria for use in the pediatric population is currently unavailable (MacAllister et al., 2012). This is partially attributable to the inadequate performance of proposed criteria in younger cohorts

(Callen, Shroff, Branson, Lotze et al., 2009). The KIDMUS (Kids with MS) criteria, for instance, which focuses on the presence of well-defined lesions and ovoid plaques, have demonstrated utility in prediction of relapse, but has been noted to lack sensitivity (Callen, Shroff, Branson, Lotze et al., 2009; Neuteboom et al., 2008). More recent criteria aimed at distinguishing pediatric MS from disseminated encephalomyelitis (ADEM) and a host of non-demyelinating neurologic disorders have been developed (Callen, Shroff, Branson, Lotze et al., 2009; Callen, Shroff, Branson, Li, et al., 2009). Higher frequencies of periventricular lesions and hypointense cerebral plaques (i.e., black holes) that serve as indicators of disease progression and demyelination are particularly relevant to the differentiation of MS and ADEM; whereas, the presence of multiple T2 lesions in the periventricular and brainstem regions are relevant to differentiating MS from non-demyelinating disorders (MacAllister et al., 2012).

Comparisons of various criteria sets indicate that the MS-ADEM criteria are most useful in distinguishing between a first MS attack and monophasic ADEM, with 75% sensitivity and 95% specificity (Ketelslegers et al., 2010). Nevertheless, pediatric MS is currently defined according to adult criteria as previously discussed, with some addendums (Chabas et al., 2011). Foremost, demyelinating events cannot meet criteria for ADEM. Second, the criteria for dissemination in time can be met with subsequent MRI evidence of new lesions as early as 3 months after the initial event. Finally, an initial event that meets criteria for ADEM can only be considered a first episode of MS when

the subsequent clinical course is characterized by a second episode with new detectable lesions or by a third event 3 months after the second episode (Chabas et al., 2011).

Differential diagnosis. In addition to evidence related to lesion dissemination in space and time, one of the fundamental features in diagnosing MS is the exclusion of other conditions with similar clinical and radiological presentations (Al Futaisi, 2007; Banwell, Ghezzi, et al., 2007). White matter abnormalities detected via MRI can be attributed to a wide-range of diagnoses (Yeh et al., 2009), including those with infectious, inflammatory, metabolic, and neurodegenerative features (Al Futaisi, 2007; Yeh et al., 2009). Advances in the operationalization of pediatric MS and related diseases such as ADEM, clinically isolated syndrome (CIS), optic neuritis (ON), transverse myelitis (TM), and neuromyelitis optica (NMO) have been instrumental in differential diagnosis (Chabas et al., 2011). While the distinction between ADEM and MS is challenging, various distinguishing features have been proposed. Monophasic ADEM requires a polysymptomatic presentation suggestive of extensive brain, brain stem, spinal cord, and optic nerve involvement in addition to encephalopathy (Hahn & Tenembaum, 2011). Course duration typically reflects fluctuating symptoms and variable MRI findings for a period of up to 3 months. While residual neurologic deficits can emerge, most symptoms appear to resolve over time (Hahn & Tenembaum, 2011).

CIS, on the other hand, can have a monofocal or polyfocal presentation, but does not typically encompass encephalopathy, except in cases of brainstem and hemispheric involvement (Chabas et al., 2011). A common example of CIS is ON, a unilateral or

bilateral inflammatory disorder affecting the optic nerve that is often considered an initial manifestation of MS (Waldman & Balcer, 2011). Like ON and ADEM, TM has a monophasic presentation that primarily targets the spinal cord, leading to sensory-motor and autonomic dysfunction (Pidcock & Sebire, 2011). Unlike MS, the clinical sequelae of TM is focused on sensory (e.g., tingling, numbness) and motor (e.g., independent ambulation) difficulties, in addition to impaired bladder control and deficits in self-care (Pidcock & Sebire, 2011; Pidcock et al., 2007). A final, and particularly debilitating disease that has similar features to MS, is neuromyelitis optica (NMO; McKeon & Lotze, 2011). This rare condition has a propensity to affect the optic nerves and spinal cord, often leading to blindness and paraplegia. Unlike MS, NMO is characterized by the rapid accrual of physical disability subsequent to spinal and optic nerve attacks (McKeon & Lotze, 2011).

Cognition and Multiple Sclerosis

Myelin and Cognition

Myelin is a compressed multilayered cell membrane produced by oligodendrocytes that serves as electrical insulation for the rapid conduction of nerve impulses throughout the CNS (Fields, 2008). Myelin increases conduction velocity, or the speed at which an electrochemical signal propagates down a neural pathway, by fundamentally altering the transmission of impulses during neuronal firing. A central concept in synaptic plasticity and learning is the temporal coincidence of firing between differentially located presynaptic neurons and postsynaptic cells. Specifically, in order to

achieve synchronous arrival of inputs, axonal conduction velocity must be delayed or accelerated, based on proximity. Time of speed is thus a critical element in information processing and synaptic function (Fields, 2008).

Myelin plays an integral role in information processing given its ability to influence conduction velocity and electrical impulse activity between distant cortical regions (Fields, 2008). Specifically, myelin regulates axonal diameter, myelin sheath thickness, and the nodal structure and molecular composition of ion channels located within the paranodal region. Proper myelin function is therefore critical because while synaptic dysfunction is the cellular basis for a large group of disorders, disruptions in functional connectivity between distant brain regions can significantly impact information processing. In disorders such as MS, for instance, the distance ions can travel during a graded potential decreases as the loss of myelin progresses. A mass reduction in available and functional sodium channels and subsequent cessation of action potentials and conduction can lead to slowing and desynchronization. Such defects in myelin insulation have been implicated in impaired cognitive functioning and disorganized thinking (Fields, 2008).

The formation and maturation of myelin involves a predictable and orderly progression from central to peripheral, caudal-to-rostral, and dorsal-to-ventral (Dietrich et al., 1988). While a vast majority of myelination occurs during infancy and early childhood, this developmental process extends well into the second decade of life and coincides with the massive restructuring of synaptic connections within the cerebral

cortex (Fields, 2008; Nagy et al., 2004). Experimental and functional imaging studies suggest the involvement of myelin in cognition, learning, memory, and skills acquisition. In fact, myelination of particular brain regions coincides with the development of specific cognitive functions, including language (e.g., reading, vocabulary) and a range of executive processes. Moreover, myelination correlates with individual variations in cognitive development, intelligence, decision speed, and working memory (Fields, 2008; Nagy et al., 2004).

Neurocognitive Impairment

Cognitive impairment in MS patients has historically been considered a less imperative disease-related consequence when compared to physical disability (Julian et al., 2011). In recent years, however, the prevalence of MS-related cognitive impairment and its impact on daily functioning and quality of life has become increasingly recognized. The cognitive deficits associated with adult-onset MS are particularly well-established and thoroughly described in the literature. Specifically, the profile of adult cognitive impairment shows prominent involvement of episodic memory, complex attention, processing speed, executive functions, and visuospatial abilities. Semantic memory, attention span, language functions, and overall intelligence, on the other hand, remain relatively intact (Amato, Zipoli, et al., 2008).

Until recently, the cognitive impact of pediatric-onset MS had been largely uncharted (Julian et al., 2011). This recognized, researchers and clinicians have long speculated that children may be particularly prone to higher degrees of cognitive

impairment in comparison to adults with comparable disease severity (Julian et al., 2011; MacAllister et al., 2012). This was largely based on the fact that the neuropathological processes of MS, including inflammation, blood-brain barrier breakdown, and demyelination co-occur with the developmental processes of myelination in the developing CNS. Thus, in addition to damaging established white matter and neuronal pathways, MS in younger cohorts interferes with the development of structural regions and developing networks that are crucial to cognition and the acquisition of skills (Fields, 2008; Julian et al., 2011; MacAllister et al., 2012).

Despite this noteworthy speculation, limited published research is available on the neuropsychological implications of pediatric MS (Julian et al., 2011). Within this context, the majority of studies evaluating cognitive impairment among children have been based on single case studies or clinical series with small samples sizes, inadequate control groups, and limited measurement techniques. Dale and colleagues (2000), for instance, published data on two pediatric MS patients who demonstrated cognitive impairment, as indicated by performance on global measures of intelligence. However, given the relative insensitivity of global intelligence tasks to the full spectrum of cognitive difficulties, the magnitude and extent of impairment in these patients was ambiguous and likely underrepresented (MacAllister et al., 2012). In 2004, McCann, Farmer, and Patel published a single case study of an 11-year-old male with MS who demonstrated significant deficits in information processing speed, visual-motor skills, and executive functioning. In addition, the young boy exhibited various mood-related symptoms and

experienced significant decline in academic performance. While this case highlighted areas for further investigation, it also demonstrated the need for case series and population studies (McCann et al., 2004). Since these publications, established pediatric MS research groups have released cognitive data on larger samples of children with MS (MacAllister et al., 2012).

Banwell and Anderson (2005) examined cognitive functioning in a series of 10 cases utilizing a neuropsychological battery including measures of global intelligence, verbal and visual memory, attention, receptive and expressive language, and executive functioning. All patients demonstrated impairment on at least one measure, with some children showing impairment on most or all areas. Specifically, memory and language deficits, as defined by scores below one standard deviation from the normative mean, were indicated. Of note, higher degrees of impairment coincided with longer disease duration. However, disease severity, as measured by the Expanded Disability Status Scale (EDSS), was low, confirming that cognitive impairments can emerge independently from disease-related physical manifestations (Banwell & Anderson, 2005).

In the United States, MacAllister and colleagues (2005) examined 37 cases using a battery designed to assess attention, language, memory, visual-spatial processes, and motor functions. Though EDSS scores indicated only mild neurologic dysfunction, significant cognitive impairment, as defined by performance falling 1.5 or more standard deviations below the normative mean on two or more tasks, was evidenced in 35% of patients. A total of 59% demonstrated impaired performance on at least one measure. The

most frequently impaired domain was complex attention, as measured by Part B of the Trail Making Test, affecting 30% of patients. Unlike studies in adult MS, expressive language (i.e., confrontation naming), as measured by the Boston Naming Test, and receptive language, as measured by the Listening to Paragraphs subtest of the Clinical Evaluation of Language Fundamentals, were impaired in a significant proportion of the cohort, affecting 19 and 14% respectively. Additionally, although only 3% demonstrated impaired verbal memory, a significant decrease in performance was evidenced on the delayed portion of the task, affecting 19% of patients. Other domains, including immediate visual memory, verbal fluency, and visual-spatial functions, remained relatively intact. Of note, several clinical variables were significantly related to performance. In particular, poor cognitive performance was associated with higher neurologic dysfunction, total number of relapses, and disease duration (MacAllister et al., 2005).

In 2008, an Italian research group published information on the largest cohort of children and adolescents with MS to date (Amato, Gorretti, et al., 2008). The study included 63 pediatric MS patients, with comparable gender representation, and a comparison group of 57 demographically-matched healthy controls. All patients demonstrated a relapsing-remitting disease course and a mean EDSS score of 1.5. The neuropsychological battery included a broad measure of intelligence (i.e., WISC-R) in addition to various tasks of verbal and visual-spatial memory, sustained attention, abstract reasoning, and expressive and receptive language (Amato, Gorretti, et al., 2008).

In comparison to matched controls, 28% of patients in the MS group demonstrated significant impairments on the verbal and performance IQ scores of the WISC-R, while 8% exhibited Full Scale IQ scores below 70, reaching the range of performance that is associated with intellectual disability (Amato, Gorretti, et al., 2008). Significant cognitive impairment, defined as failure (i.e., at or below fifth percentile of control performance) on 3 or more tests, was present in 31% of this cohort. An additional 22% exhibited minor degrees of cognitive dysfunction, as defined by impairment on at least two measures. Similar to previous findings, verbal comprehension, visual-spatial memory, complex attention, and executive function were among the most frequently impaired domains. Moreover, semantic and phonemic verbal fluencies were compromised in 22 and 17% of the cases, respectively. Of note, younger age at onset and IQ scores less than 90 were the only significant predictors of cognitive dysfunction (Amato, Gorretti, et al., 2008). In a more recent study (Portaccio et al., 2009), the same Italian group identified cognitive impairment in 41% of children and adolescents with MS utilizing failure (i.e., at or below fifth percentile of control performance) on 4 or more measures as the threshold. The mean number of impaired tests in MS patients was 4.4 tests compared to 1 test in the healthy control group. In comparison to controls, MS patients were found to perform significantly worse on tasks of verbal and nonverbal abilities, episodic memory, executive functioning, attention, and language (Portaccio et al., 2009).

Overall, it is estimated that at least one-third of children and adolescents with MS exhibit significant cognitive impairment (Julian et al., 2011). Current research suggests that the profile of cognitive dysfunction in pediatric-onset MS resembles, to some extent, the known profile of adult MS. Specifically, involvement of memory functions, complex attention, processing speed, and various aspects of executive functioning appear to be implicated in both. These findings are expected given the critical role of myelin and the relevance of neuronal connectivity across these processes. While studies on children and adolescents with MS are limited, it is important to highlight that the breadth of deficits in pediatric-onset MS is likely to be more extensive, impacting domains that remain relatively intact in adults. It has been hypothesized, for instance, that language faculties are particularly vulnerable in younger patients who sustain disruption during the critical periods of linguistic development. There is, in fact, some evidence to suggest that in addition to demonstrating deficits in verbal fluency, children and adolescents struggle on tasks of conformation naming and receptive language. Another noteworthy finding in this age range is the impact of MS on overall intellectual functioning. While the assessment of processing speed and other processes known to be compromised in MS are integral to the assessment of overall intellectual abilities in children, performance in younger cohorts suggests some relationship between diffuse neuropathology in the CNS and a global impact on the development of intellectual abilities (Julian et al., 2011).

Similar to cases of adult-onset MS, the relationship between cognitive dysfunction and various primary clinical features, including disease duration, disability

scores, and relapse frequency, remains elusive in pediatric-onset cases (Julian et al., 2011). While increasing disability levels and disease duration appear to be associated with cognitive impairment, deficits across various domains are frequently observed in patients without overt physical signs of disability (Julian et al., 2011).

Longitudinal Outcomes

Longitudinal studies of adults with MS have generally demonstrated some degree of cognitive loss across time (Julian et al., 2011). A 3-year study (Kujala, Portin, & Ruutiainen, 1997) examining cognitive functioning in a cohort of 42 adults with MS, for instance, found evidence of progressive cognitive decline in patients that were cognitively intact at baseline, in addition to support for more pronounced cognitive loss in patients who initially presented with some degree of cognitive impairment. Progressive deterioration was primarily observed on tests of memory, learning, attention, and visuo-motor processes, which have all been reported to demonstrate some sensitivity to MS (Kujala et al., 1997). A similar 10-year study (Amato, Ponziani, Siracusa, & Sorbi, 2001) that included a cohort of 45 MS patients in the incipient stage of the disease revealed declines in cognitive capacities among 63% of patients over time. Cognitive declines were primarily associated with verbal memory, abstract reasoning, linguistic processes, attention, and short-term spatial memory. Multiple regression analysis revealed that a secondary progressive disease course, higher degrees of physical disability, and increasing age were independently associated with the extent of cognitive decline (Amato et al., 2001).

While recent research endeavors have attempted to elucidate the cognitive impact of pediatric-onset MS, there is a dearth of data related to the progression of cognitive outcomes over time (Amato et al., 2010; MacAllister, Christodoulou, Milazzo, & Krupp, 2007). To date, only two longitudinal studies exploring the neuropsychological outcomes among pediatric cohorts have been published (Julian et al., 2011). These studies have served as preliminary, yet integral, components of the pediatric literature on MS, and foundational predecessors to the current study.

In the only study of its kind within the United States, MacAllister, Christodoulou, and colleagues (2007) recruited a group of 12 children with MS under the age of 17 years and administered a battery of tests aimed at surveying a wide-range of cognitive domains including: attention, executive functioning (cognitive flexibility), expressive language (verbal fluency), receptive language, verbal and visual learning and memory, and visual motor integration. Cognitive impairment on each subtest was indicated by a score of 1.5 or more standard deviations below the mean. Of note, time between baseline and follow-up averaged 21.58 months. At baseline, 10 of the 12 participants demonstrated impaired performance on at least one task. While the same results were reproduced at follow-up, it is important to note that these were not necessarily the same individuals from baseline. One participant who had demonstrated impairment across three tasks at baseline, for instance, was not classified as impaired on any of the tasks at follow-up. On the other hand, another participant who did not demonstrated impairment at baseline, had scores indicative of impairment on two tasks upon re-evaluation. Overall, significant

increases in impairment were evidenced across tasks of executive functioning, visual motor integration, verbal fluency, and visual learning. Of note, part B of the Trail Making Test, a measure of complex attention and executive functioning, demonstrated the largest increase in impairment frequency (MacAllister, Christodoulou, et al., 2007).

To further investigate the evolution of neurocognitive functioning among children with MS, Amato and colleagues (2010) assessed a group of 56 children utilizing a battery comprised of measures that addressed the following domains: visual and verbal learning and memory, complex attention, executive functioning (planning), expressive language (verbal fluency), and receptive language. Cognitive impairment was defined as failure (i.e., performance below the 5th percentile) on at least three tests. Assessment results at the 2-year follow-up revealed significant adverse outcomes, with 70% of patients fulfilling the cognitive impairment criteria, as compared to 31% at baseline. An additional 6.5% of patients (i.e., 76.5% total) were predicted to exhibit deteriorating cognition on the basis of cognitive change index scores. Results were indicative of a 17% decrease in cognitive function from the time of baseline. Changes were particularly prominent in tasks of verbal memory, complex attention, verbal fluency, and receptive language (Amato et al., 2010).

Despite noted efforts, the literature with respect to cognitive prognosis in pediatric MS patients remains extremely limited. Existent longitudinal studies have primarily utilized a frequency-based approach emphasizing the number of patients performing in the impaired range or the number of failed tests at follow-up relative to baseline. Of note,

time between baseline and follow-up has been relatively similar (21.58 to 24 months) across studies. Results from these studies suggest that pediatric MS patients exhibit significant cognitive declines in verbal learning and memory, complex attention, and verbal fluency within a relatively short period of time. However, to better understand the evolution and impact of MS on cognitive functioning and the domain-specific rates of decline, it will be crucial to evaluate pediatric MS patients across varying time points. Thus far, studies have failed to capture results within a 12 to 18 month period or after a 24 month period. Equally important is the continued exploration of the relationship between cognitive impairment and other clinical and disease-related factors.

Assessment of Neurocognitive Functioning in Multiple Sclerosis

Neuropsychological Batteries

The development of systematic approaches by which to perform routine cognitive evaluations has become an increasingly imperative task, given the noted prevalence of cognitive impairment across adult and pediatric MS cohorts (Julian et al., 2011). Despite a large consensus regarding the importance of early cognitive evaluations, uniform procedures and methods remain largely at bay (Possa, 2010). In fact, the utility of extensive versus brief cognitive batteries has become an important issue as it relates to exploring the neuropsychological functioning of patients with MS. Extensive batteries provide the opportunity to explore the domains that appear to be more pathologically involved in the disease as well as overall cognitive functioning (Possa, 2010). Nevertheless, such batteries can be time-consuming and impractical ways by which to

ascertain data (Portaccio et al., 2009). Specifically, not only can they present obstacles for patients experiencing fatigue, but for practitioners charged with the task of evaluating multiple patients within a given period of time (Possa, 2010). As such, it has been increasingly suggested that patients undergo a brief battery comprised of various diagnostic tools with demonstrated sensitivity to cognitive changes across domains that are particularly impacted by MS (i.e., working memory, attention, processing speed, executive functions; Possa, 2010).

Given the heterogeneity of cerebral pathology in MS, devising a single, brief measure that is sufficiently sensitive to the cognitive impairment evidenced in this disease is not a likely feat (Benedict & Zivadinov, 2007; Strober, Englert, Munschauer, Weinstock-Guttman, Rao, & Benedict, 2009). Nonetheless, researchers have identified two brief batteries that are useful in monitoring cognitive functions in adults with MS: the Rao Brief Repeatable Neuropsychological Battery (BRNB) and the Minimal Assessment of Cognitive Function in MS (MACFIMS). Both batteries are purported to have solid validity and to have reached a threshold of wide acceptance and replicability. As such, these batteries are frequently utilized in the contexts of research and clinical practice (Boringa et al., 2001; Strober et al., 2009; Yeh, Parrish, & Weinstock-Guttman, 2011).

The BRNB was developed by the Cognitive Function Study Group of the National MS Society (Boringa et al., 2001; Solari, Mancuso, Motta, Mendozzi, & Serrati, 2002; Strober et al., 2009). Measures were empirically chosen by identifying and extracting the most sensitive tasks from a comprehensive set of 23 tests that were initially

chosen according to guidelines for Neuropsychological Research in MS. The BRNB encompasses five tests, including measures of sustained attention, concentration, and speed of information processing (i.e., Paced Auditory Serial Addition Test [PASAT], Symbol Digit Modalities Test [SDMT]), verbal learning and delayed recall (i.e., Selective Reminding Test [SRT]), visuospatial learning and delayed recall (i.e., 10/36 Spatial Recall Test), and verbal fluency (i.e., California Oral Word Association Test [COWAT]; Boringa et al., 2001; Solari, 2002; Strober et al., 2009).

In 2001, an expert panel of neurologists, neuropsychologists, and psychologists convened to establish the most reliable and valid brief neuropsychological measure for use in adult MS patients to date (Strober et al., 2009). This consensus conference resulted in the emergence of the MACFIMS. While several of the elements from the BRNB were preserved, the MACFIMS employed different tasks of verbal and visuospatial learning and delayed recall. Additionally, measures of executive and visuospatial functions were incorporated. These substitutions and additions were based largely on increasing knowledge with regard to the psychometric properties of tests in addition to evolving research on commonly impaired domains in adult MS (Strober et al., 2009).

Specifically, the MACFIMS comprises seven tests that assess processing speed, working memory, and attention (i.e., PASAT, SDMT), verbal and visuospatial learning and memory (i.e., California Verbal Learning Test-Second Edition [CVLT-II], Brief Visuospatial Memory Test-Revised), verbal fluency (i.e., COWAT), visuospatial processing (i.e., Judgment of Line Orientation Test), and executive functions

(i.e., Delis-Kaplan Executive Function System Sorting Test; Benedict et al., 2006; Strober et al., 2009). In a recent study comparing the relative sensitivity and utility of these batteries, Strober and colleagues (2009), concluded that the BRNB and the MACFIMS exhibited similar discriminative validity, as they equally distinguished MS patients from healthy controls, with 79 and 83% accuracy, respectively. Of note, the immediate and delayed trials of the SRT and the CVLT-II were comparable in sensitivity, specificity, and predictability. The CVLT-II, however, is more widely used and has a more well-established test-retest reliability (Strober et al., 2009). Despite such differences, both batteries are frequently used across adult MS samples.

As is the case in adult MS cohorts, a proliferation of literature on the use of brief batteries with pediatric MS patients has recently begun to emerge. The Brief Neuropsychological Battery for Children (BNBC), proposed by Portaccio and colleagues (2009), for instance, encompasses four measures that tap into verbal knowledge, processing speed, executive functioning, and verbal memory. Specifically, the Vocabulary task from the Wechsler Intelligence Scale for Children-Revised (WISC-R), the Symbol Digit Modalities Test (SDMT), parts A and B of the Trail Making Test (TMT A and TMT B), and the Selective Reminding Test (SRT and SRT-Delayed) are incorporated. To assess its ability to detect cognitive impairment, the BNBC was tested against a more comprehensive set of measures. A fairly liberal cut-off of one or more tests with impaired scores, defined as performance below the 5th percentile relative to controls, was utilized. The sensitivity and specificity of the BNBC was found to reach 96

and 81%, respectively. These results indicate a reasonable level of cognitive impairment detectability and clinical utility (Portaccio et al., 2009).

The Pediatric MS Centers of Excellence sponsored by the National Multiple Sclerosis Society have specified the use of a research-oriented neuropsychological battery comprised of various core measures noted to exhibit sensitivity to various impairments commonly seen in pediatric MS (Julian et al., 2011). Impairment areas assessed by the Neuropsychological Battery for Pediatric MS (NBPMS) include: an overall estimate of ability, verbal learning and memory, attention, working memory, processing speed, visual-spatial processing, and executive functioning. Specifically, this battery includes the Vocabulary, Similarities, Block Design, and Matrix Reasoning tasks from the Wechsler Abbreviated Scales of Intelligence (WASI), the California Verbal Learning Test (CVLT-II and CVLT-C), the Contingency Naming Test, and either the Coding and Digit Span tasks from the Wechsler Intelligence Scales for Children, Fourth Edition (WISC-IV) or the Digit Symbol and Digit Span tasks from the Wechsler Adult Intelligence Scales (WAIS-III), depending on age. Of note, the NBPMS is more time-consuming than the BNBC, and unfortunately, no studies investigating the utility of this battery for detecting cognitive impairment in pediatric MS have been published (Julian et al., 2011).

Information pertaining to the relevance and structure of brief neuropsychological screeners has also been gleaned from other pediatric populations, including cancer. Krull and colleagues (2008), for instance, developed and validated a brief neurocognitive

battery, the DIVERGT, to overcome similar barriers (e.g., costs, time restraints, limited access to professional resources) encountered in the assessment of patients with MS. This brief battery is comprised of performance-based measures specifically designed to assess childhood cancer and treatment-sensitive domains. Specifically, the battery includes the Digit Span task from the third edition of the WISC or WAIS, the Verbal Fluency task from the Multilingual Aphasia Examination (MAE), the Grooved Pegboard Test, and parts A and B of the Trail Making Test from the Halstead-Reitan Neuropsychological Test Battery (Krull et al., 2008).

Test-retest reliability, as assessed by comparing the performance of 48 patients at baseline and follow-up, was noted to be good, with an overall $r = 0.72$ (Krull et al., 2008). Equally sound discriminative validity, as evidenced by significant differences in the performance of patients at high risk versus low risk for neurocognitive problems, was also reported. To determine the sensitivity and specificity of the battery, 52 patients who had completed the screener were given a comprehensive follow-up assessment for performance comparisons. Sensitivity and specificity of the DIVERGT were 94 and 63%, respectively, for impaired global intellect, 87 and 57%, respectively, for reading impairment, and 100 and 50%, respectively, for mathematical impairment. Overall, the DIVERGT was a better predictor of global cognitive and academic functioning in comparison to specific parent reports (Krull et al., 2008).

Although these batteries seem promising, it is commonly acknowledged that research related to cognitive dysfunction in children with MS is in its infancy (Yeh et al.,

2011). As such, a predominant number of neuropsychological batteries are structured to assess a comprehensive set of cognitive domains, including processing speed, learning and memory, language, attention and executive functioning, visual-spatial processing, and sensorimotor functioning. Specific focus has been placed on areas most commonly purported to be affected in adults with MS, such as working memory, processing speed, and visual-spatial processing. While general measures of cognitive functioning and academic achievement have been included (Yeh et al., 2011), time constraints coupled with a more prominent focus on specific cognitive domains has supported their exclusion (MacAllister et al., 2005). Measures that tap into a child's psychosocial functioning are also common, given the noted concerns in the areas of internalizing behaviors (e.g., depression, anxiety), fatigue, and peer relationships (Yeh et al., 2011). While most pediatric MS clinics and centers employ a comparable standardized battery, differences in the specific measures used are expected (Yeh et al., 2011).

Neuropsychological Measures in Pediatric Multiple Sclerosis

Despite expected variability across assessment batteries, neuropsychological tests selected for use with pediatric MS patients should be grounded on specific criteria (Till et al., 2011). Foremost, batteries should be comprised of measures intentionally developed to serve as standardized assessment tools. Additionally, measures with adequate normative data encompassing a broad age range, minimal floor and ceiling effects, and good reliability and validity are crucial. Lastly, it is critical that selected measures assess those cognitive abilities that have been previously reported to be impacted in patients

with childhood MS, including but not limited to: attention, processing speed, language, learning and memory, visual-spatial processing, and executive functioning (Till et al., 2011). To this effect, various measures, and their utility with pediatric MS patients, have been discussed in the literature.

The Symbol Digit Modalities Test (SDMT) assesses processing speed in the visual-spatial domain without the incorporation of a motoric output demand (Smerbeck et al., 2011). To establish the sensitivity and validity of this measure in pediatric MS, Smerbeck and colleagues (2011) compared the performance of 43 MS patients and 45 controls on the SDMT and the Brief Visuospatial Memory Test-Revised (BVMTR). The MS patients perform significantly worse than their demographically matched controls on both the BVMTR and SDMT. Additionally, both measures yielded significantly weak performance in MS patients, although the effect size for the BVMTR (d = .87) was found to be larger than that of the SDMT (d = .69; Smerbeck et al., 2011). Similar findings had been reported in the study conducted by Amato, Gorretti, and colleagues (2008), in which the SDMT was noted to represent an area of cognitive impairment in 28 to 36.8% of MS patients at baseline and follow-up, respectively (Amato, Gorretti, et al., 2008). These findings were upheld in a more recent quasi-experimental study (Till et al., 2011), wherein a greater proportion of patient with MS scored in the below average range on the SDMT relative to controls (12% versus 0%). Given differentials in performance between affected and healthy groups across studies, it can be concluded that the SDMT is a useful tool in the assessment of children with MS (Till et al., 2011).

The Trail Making Test (TMT) is a multifactorial task that is divided into two parts. Part A (TMT A) requires rapid visual scanning and motor speed; whereas, part B (TMT B) incorporates the more complex skill of shifting attention (MacAllister, Christodoulou, et al., 2007). Given its high sensitivity to brain dysfunction, the TMT is one of the most frequently utilized neuropsychological measures (Baron, 2004; MacAllister, Christodoulou, et al., 2007). Its utility in the assessment of neuropsychological functioning in children with MS has been derived from various sources. In a previously discussed study conducted by Amato, Gorretti and colleagues (2008), for instance, the frequency at which patients demonstrated performance in the impaired range was noted to increase by task, with both aspects of the TMT showing the largest increase in impairment frequency (50%) of study participants (Amato, Gorretti, et al., 2008). A 2-year follow-up on the same cohort indicated increased impairments on the TMT A (57.9%; Amato et al., 2010).

A similar longitudinal study conducted by McAllister, Christodoulou, and colleagues (2007) indicated that the TMT B was one of the most frequently impaired tasks at baseline, with the TMT A closely following. At a 2-year follow-up evaluation, the frequency at which patients performed in the impaired range rose on several tasks, with the TMT B showing the largest increase in impairment frequency (MacAllister, Christodoulou, et al., 2007). Lastly, in a more recent study (Till et al., 2011), a pediatric MS cohort was noted to perform significantly below average on the TMT A relative to

healthy controls. Taken together, these results appear to demonstrate the utility and clinical sensitivity of this measure as it relates to pediatric MS.

The WISC-IV and the WAIS-III are members of a widely used series of intellectual assessments, with well-documented utility across populations (Wechsler, 1997; 2003). While the Wechsler tests have been primarily used to obtain estimates of general intellectual functioning in pediatric MS patients, the inclusion and sensitivity of the processing speed and working memory tasks have been evidenced, to some degree, in pediatric MS studies. In a study conducted by Holland, Graves, Greenberg, and Harder (2012), for instance, approximately 8% of participants were noted to exhibit clinically significant impairment, defined as performance at or below 1.5 standard deviations from the normative mean on the Digit Span task. A cross-sectional, multi-center study, however, revealed higher rates of impairment on the Digit Span task (27%) and the Coding task (35%; Julian et al., 2013). While additional investigations utilizing the working memory and processing speed tasks from the Wechsler tests are warranted, these results provide preliminary evidence of their utility.

The Beery-Buktenica Developmental Test of Visual-Motor Integration (Beery VMI) is a measure of motor coordination and spatial organization that requires an individual to copy designs of increasing complexity (Beery & Beery, 2006). Two studies incorporating the Beery VMI have demonstrated its utility in pediatric MS evaluations. In the previously discussed longitudinal study by McAllister, Christodoulou, and colleagues (2007), the Beery VMI was noted to be the second most problematic measure after a

2-year period, with impairment rates of 33% at follow-up compared to 8% at baseline (MacAllister, Christodoulou, et al., 2007). Similar rates (34%) of mild impairment, defined as performance below one standard deviation from the mean, were also reported in a more recent study (Julian et al., 2013). This latter study also noted that 16% of patients exhibited moderate to severe impairment (i.e., scores falling at least 2 standard deviations below the mean) on this measure. While the incorporation of the Beery VMI in pediatric MS studies has been limited, preliminary data indicate that this measure may be useful in assessing this population.

Although studied to a lesser extent, the grooved pegboard, the Delis-Kaplan Executive Function System (D-KEFS) Verbal Fluency Test, and the CVLT tests have also been noted as having sensitivity to pediatric MS. The grooved pegboard is a measure of fine motor speed, hand-eye coordination, and manual dexterity (GPT; Klove, 1963). While motor functions are commonly assessed in pediatric MS patients, one study (Julian et al., 2013), has demonstrated the sensitivity of the grooved pegboard with pediatric MS patients. In this study, the grooved pegboard was reported as the test with the most frequent rate of impairment. Specifically, 20% of patients exhibited mild impairment when using their dominant hand; whereas, 23% exhibited moderate to severe impairment. The frequency of combined impairment rose to 57% when patients used their non-dominant hand. The authors of the same study also reported a combined impairment frequency of 29% on the CVLT-C, a measure of verbal learning and memory that encompass 5 learning trials, an interference trial, and a recognition task (Delis, Kramer,

Kaplan, & Ober, 1994). Lastly, in the previously mentioned study by Holland and colleagues (2012), 19.2% of participants exhibited clinically significant impartment on the D-KEFS Verbal Fluency Test (Holland et al., 2012). While additional studies are needed to further establish the utility of the aforementioned measures, available data suggests that these measures may be useful in detecting domain-specific cognitive dysfunction in pediatric MS patients.

Factors Affecting Assessment Performance

Psychological distress. The prevalence of psychological distress caused by affective comorbidities among adults with MS has been extensively explored. Clinical depression, for instance, can affect up to 50 to 70% of patients with MS over the course of their lifetime and has been associated with a host of adverse outcomes (Venkateswaran, Bennett, & Ness, 2011), including cognitive impairment, suicidal ideation and completed attempts, diminished relationships, and non-compliance to disease modifying treatments. Undoubtedly, depression is regarded by adult MS patients as one of the primary determinants of quality of life (Feinstein, 2011). Anxiety, although studied to a lesser extent, can affect between 23.5 and 41% of adults with MS and has been associated with suicidal ideation and self-harm. While less is understood about the psychiatric well-being of pediatric patients (Julian et al., 2011), qualitative investigations (e.g., Boyd & MacMillan, 2005) exploring themes related to emotional distress in pediatric cohorts have revealed significant levels of stress related to treatment adherence, unpredictability of relapse and symptoms, and school-related difficulties. Quantitative

studies evaluating the psychological characteristics of children and adolescent with MS have been largely inconclusive, with rates of psychological distress ranging from very rare to over 50% (Julian et al., 2011). Nevertheless, there is increasing evidence to suggest comparable rates of depression and anxiety between pediatric and adult cohorts (Amato, Zipoli, et al., 2008; MacAllister, Boyd, Holland, Milazzo, & Krupp, 2007).

While studies exploring the relationship between psychological distress and cognition among pediatric MS cohorts are largely nonexistent, a recent study by Holland and colleagues (2012) found a negative correlation between depression, anxiety, and executive functioning. Specifically, higher degrees of parent-reported depression and anxiety were associated with lower performance scores on tasks of working memory and cognitive flexibility (Holland et al., 2012). Further implications of psychological distress on cognition can be gleaned from the adult literature (Julian et al., 2011). Specifically, depression has been known to affect planning ability, working memory capacity, and attentional processes, while anxiety has been implicated in executive functioning among adults with MS (Julian & Arnett, 2009; Julian et al., 2011). The relationships between psychological distress and cognitive impairment in adults are multifaceted and interactive. Cognitive impairment, for instance, may be a functional consequence of depressive symptomology (e.g., reduced attention and concentration); alternatively, both depression and cognitive impairment may stem from damage to similar neuropathological pathways (Julian et al., 2011). Relationships between lesions, depression, and cognition, for instance, have been identified in MS patients, and the presence of cognitive

impairment has been associated with poor response to conventional antidepressant interventions (Julian & Mohr, 2006; Zorzon et al., 2002).

Fatigue. MS-related fatigue is characterized by an overwhelming sense of lethargy coupled with a profound decrease in energy (Julian et al., 2011; Venkateswaran et al., 2011). The pathophysiology of MS-related fatigue is multifactorial and complex, involving dysregulation of the immune system, alterations in the CNS, and changes to neuroendocrine and neurotransmitter processes (MacAllister & Krupp, 2005). Secondary factors including heat exacerbation, excessive physical exertion, sleep problems, and mood disorders play contributing roles (Julian et al., 2011; MacAllister & Krupp, 2005; Venkateswaran et al., 2011). Fatigue is the most commonly experienced and debilitating symptom in adults with MS, significantly impacting daily functioning, productivity, and overall quality of life (Julian et al., 2011; MacAllister et al., 2009). Unfortunately, research related to the incidence and severity of fatigue among pediatric cohorts is less established; nevertheless, recent studies have reported rates between 40 and 50% (Amato, Zipoli, et al., 2008; MacAllister et al., 2009). Though these rates are slightly lower than those evidenced in the adult literature, the fluctuating and disruptive nature of fatigue has certainly been associated with decrements in cognitive functioning (Julian et al., 2011).

Cognitive fatigue has been defined as a decline in cognitive performance over a brief period of time that can emerge in the absence of participation in physically demanding activities and observable physical impairments (Julian et al., 2011; MacAllister et al., 2009). In a study conducted by Krupp and Elkins (2000), for instance,

patients with MS demonstrated apparent declines in performance on measures of executive functioning and verbal memory over a 4-hour session. In contrast, healthy controls demonstrated improvements on these measures across time. While practice effects likely contributed to performance enhancement in healthy controls, these findings allude to the relevance of brief screenings and testing over multiple sessions in MS patients (Julian et al., 2011; Krupp & Elkins, 2000).

Steroid treatment. As a result of their ability to modulate the immune system, corticosteroids are widely used to treat inflammatory conditions in adults and children (Aaen, 2011). Corticosteroids are considered particularly efficacious in the treatment of acute MS relapses, given their capacity to secrete anti-inflammatory cytokines, reduce demyelination, prevent oligodendorocytic apoptosis, and inhibit axonal loss. Current practice for treating moderate to severe exacerbations in children consists of 20 to 30 mg/kg of methylprednisolone, which is typically administered intravenously on a daily basis for 3 to 5 days, with a maximum dosage of 1000 mg (Kuntz et al., 2010). While this regimen is usually followed by an oral prednisone taper, there are no data in children to suggest that tapering is beneficial. In fact, while the overall treatment regimen is common, it is largely based on adult data, as clinical trials have not been performed to evaluate their efficacy in children (Aaen, 2011).

Although corticosteroids are a leading treatment option, they are often associated with adverse physiological and neuropsychiatric side effects (Warrington & Bostwick, 2006). Physiological side effects include insomnia, osteoporosis, hyperglycemia,

hypertension, acne, and weight gain, although these effects are more highly associated with prolonged, rather than short, high dosage regimens, which are typically used to treat MS exacerbations. Neuropsychiatric side effects include mood lability, anxiety, behavioral disturbances, cognitive impairments, and psychosis (Aaen, 2011; Warrington & Bostwick, 2006). Cognitive deficits, including declarative and verbal memory deficits have been well-documented in adult populations. Deficits during short term corticosteroid therapy are consistent with hippocampal dysfunction and reversible atrophy of hippocampal neurons (Aaen, 2011; Lupien & McEwan, 1997; Newcomer, Craft, Hershey, Askins, & Bardgett, 1994). Declarative memory deficits can emerge after just 4 to 5 days of prednisone therapy. These side effects appear to be dose-dependent and reversible upon discontinuation (Newcomer et al., 1994).

Neurological impairment and ambulation. Over the last 40 years, many clinical scales have been proposed to assess neurological impairment in MS patients (Sharrack, Richard, Soudain, & Dunn, 1999). These scales are largely based on the functional assessment of several physiological systems, including pyramidal (ambulation), cerebellar (coordination and balance), brainstem (speech and swallowing), sensory (touch and pain), bowel and bladder, visual, and mental (Ghezzi & Banwell, 2011). Despite noted limitations and criticisms (i.e. inadequate ability to account for cognitive disability, overemphasis on motor performance), the EDSS is the most widely used scale for assessing neurological impairment. The EDSS quantifies disability based on a score ranging from 0 (normal neurological examination) to 10 (death due to MS; Ghezzi &

Banwell, 2011). Investigations assessing the EDSS in prospective pediatric MS cohorts have revealed median EDSS scores of 1.3 in the first year of disease and 1.5 after three years of disease. In addition to the EDSS, the Scripps Neurological Rating Scale (SNRS) and the Ambulation Index (AI) have been commonly used in adult MS research (Sharrack et al., 1999). Unfortunately, their incorporation in pediatric studies is limited.

While substantial interindividual variation with regard to neurological impairment has been documented, patients with childhood-onset MS take an average of 10 years longer than patients with adult-onset MS to reach comparable milestones of disability, including irreversible loss of ambulation (Yeh et al., 2009). That said, patients with pediatric-onset MS reach higher disability scores at a younger age, thereby challenging the generally accepted notion of a more favorable prognosis in the pediatric group (Ghezzi & Banwell, 2011; Yeh et al., 2009). Although neurological impairment and ambulation are critical components of MS, the extent to which these factors correlate with cognitive dysfunction is unclear. Specifically, while cognitive impairment can exist in the absence of physical disability, studies examining the relationship between cognitive dysfunction and neurological impairment have produced contradictory findings (Ghezzi & Banwell, 2011).

In 2005, for instance, MacAllister and colleagues found a significant relationship between EDSS scores and cognitive performance. Specifically, EDSS scores accounted for 33% of the variance in cognitive functioning. Of note, this relationship was present after controlling for dominant hand motor functioning. Moreover, a hierarchical multiple

regression revealed that EDSS performance was the strongest predictor of cognition, as compared to other clinical factors, including relapse rate, age at onset, fatigue, and disease length (MacAllister et al., 2005). A logistic regression analysis conducted by Amato, Goretti, and colleagues (2008), on the other hand, found an IQ score lower than 90 to be the best predictor of cognitive functioning, as compared to the EDSS and a host of clinical variables (Amato, Goretti, et al., 2008). Similar discrepancies have emerged in longitudinal investigations. In the study conducted by MacAllister, Christodoulou, and colleagues (2007), for instance, a statistically significant relationship between EDSS scores and number of tests in the impaired range was noted; whereas, no significant association between cognitive outcomes and EDSS scores were found in the study conducted by Amato, Goretti, and colleagues (2010). It is important to note that the samples from these studies had similar mean EDSS scores.

Summary

The proliferation of sclerotic lesions ensuing from episodes of inflammation and demyelination affects nerve transmission by slowing vital electrochemical conduction activity in MS patients. These pathogenic abnormalities result in the clinical manifestation of various physical, cognitive, and emotional symptoms (Arrondo et al., 2009). The varied constellation of symptoms associated with MS can be functionally impairing, often impacting educational performance, employment, and overall quality of life (Ghezzi & Banwell, 2011). Common cognitive consequences include deficits in learning and memory, processing speed, complex attention, and several aspects of

executive functioning (e.g., conceptual reasoning; Amato, Zipoli, et al., 2008). Other clinical symptoms include fatigue, depression, and physical impairments (e.g., ambulation).

While considerable research has been devoted to understanding the nature, prevalence, and severity of such symptoms in adult MS patients, the literature with regard to the cognitive and clinical sequelae of pediatric MS remains limited. This void can be largely attributed to the fact that pediatric MS has been the object of serious and comprehensive investigation for only a decade. Additionally, this paucity has been fueled by an almost exclusive emphasis on adult disease burden, the use of single case studies and small clinical samples, and the lack of variability in measurement techniques (Julian et al., 2011). In 2005, Banwell and Anderson published the first cohort-based study to address the need for data pertaining to the neuropsychological implications of pediatric MS. Since their publication, only a handful of cross-sectional studies (i.e., Amato et al., 2008; MacAllister et al., 2005; Portaccio et al., 2009) have emerged.

While these research endeavors have attempted to elucidate the cognitive impact of pediatric-onset MS, there is a dearth of data related to the progression of cognitive outcomes over time (Amato et al., 2010; MacAllister, Christodoulou, Milazzo, & Krupp, 2007). To date, only two longitudinal studies exploring the evolution of cognitive functioning among pediatric cohorts have been published (Julian et al., 2011). Despite the integral role that these studies have played in unraveling the long-term effects of pediatric MS, they have been limited in their scope and methodologies. Specifically, both studies

have used a similar frequency-based methodology for analyzing data as well as a comparable re-test interval. Moreover, these studies have provided only a limited amount of information with regard to the relationship between cognitive dysfunction and various clinical variables, including age at onset, disease duration, and relapse rate. Lastly, the contribution of depression and fatigue, two of the most commonly experienced and debilitating symptoms associated with MS, on the cognitive deficits experienced by pediatric MS patients has been largely overlooked.

Given the limitations in the literature, the present study assessed cognitive functioning in a cohort of children and adolescents with MS after a mean period of 14.65 months from baseline using a comprehensive battery of measures. The study also examined the relationship between cognition and various clinical factors, including age at onset, relapse rate, disease duration, fatigue, and depression. Particularly, the study explored whether, and to what extent, changes in cognition correlated with these factors. Of note, the study reported physical disability estimates based on the AI, rather than the commonly used, though controversial, EDSS. It is hoped that findings will increase the generalizability of longitudinal data, inform educational and treatment planning, and help further delineate the cognitive domains most vulnerable to disruption during active neurodevelopment.

CHAPTER III

METHOD

This chapter outlines a research study designed to elucidate the neuropsychological outcomes of pediatric MS over a mean period of 14.65 months. The primary purpose of the study was to analyze changes in performance on various measures designed to assess those neurocognitive processes with purported susceptibility to MS (e.g., processing speed, attention, working memory) in a pediatric sample. Additionally, the study explored whether, and to what extent, a relationship exists between changes in cognition and various clinical factors, including age at onset, disease duration, relapse rate, depression, and fatigue. This chapter outlines information regarding participants, procedures, measures, and data analysis techniques.

Research Participants

Data for this study was culled from a clinical sample of pediatric patients with written informed consent for participation in a retrospective cohort project approved by the Institutional Review Board (IRB) at the University of Texas Southwestern Medical Center. Participants included consecutive pediatric patients who completed serial brief neuropsychological screening evaluations as part of their routine clinical care at the Children's Medical Center Dallas Pediatric Demyelinating Diseases Clinic. Eligibility criteria included a neurologist-confirmed diagnosis of childhood MS (relapsing-remitting subtype) using revised McDonald criteria.

While most participants had a diagnosis of relapsing-remitting MS (RRMS) at the time of initial evaluation, a few had progressed from a diagnosis of clinically isolated syndrome (CIS) between baseline and follow-up. Given the effects of exacerbations and steroids on cognition, patients were excluded from the study if they were within 30 days from acute symptoms and/or steroid treatment. Additionally, patients older than 18 years of age at time of diagnosis and those with neurological conditions unrelated to MS were disqualified. Given these parameters and the rarity of pediatric MS, a final sample size of 20 participants between the ages of 5 and 16 years at time of diagnosis and 9 to 18 years at baseline were included in the study.

Procedure

As part of their routine multidisciplinary clinic visit, each participant completed a brief neuropsychological screening battery at baseline and approximately 12 to 18 months later. The test battery was administered by a trained examiner in a single 60-minute session, although test duration varied, to some extent, based on individual patient factors (e.g., age, time to complete task, level). Additionally, parents/caregivers completed questionnaires related to their child's psychological and behavioral functioning. Both parents and patients completed inventories to assess each patient's levels of fatigue and quality of life. Clinical information including diagnosis, age at symptom onset, medication use, disease duration, number of relapses, and ambulation was also collected.

To ensure confidentiality of data, each case was coded and separated from their respective case file. Cases were analyzed based on demographic information, clinical information, and specific assessment criteria in order to determine their utility and relevance to the present study. Specifically, cases with complete clinical information and subtest scores associated with specific measures of cognition were analyzed. Based upon the use of archival data and the inability to manipulate independent variables, a correlational, ex post facto research design was implemented in the present study.

Measures

Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV) and Wechsler Adult Intelligence Scale-Third Edition (WAIS-III)

Overview. The Wechsler series of intellectual assessments consists of the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV; Wechsler, 2008), the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV; Wechsler, 2003), and the Wechsler Preschool and Primary Scale of Intelligence-Fourth Edition (WPPSI-IV; Wechsler, 2012). While this series represents the most recent progression of the Wechsler scales, selected subtests from the WISC-IV and the predecessor of the WAIS-IV (i.e., WAIS-III; Wechsler, 1997) were utilized in the present study. The WAIS-III, as opposed to the WAIS-IV, was included in the study given the utilization of archival data and the discrepancy between the date the WAIS-IV was released and the date that data collection began.

The WISC-IV is a comprehensive and individually administered instrument for assessing the intellectual functioning of children between 6 and 16 years of age, whereas the WAIS-III is intended for use with individuals between the ages of 16 and 89 years of age (Wechsler, 1997; 2003). Both measures are grounded in a series of intellectual assessments that have been continuously updated over the last 60 years to reflect the needs of contemporary clinical practice. Additionally, the WISC-IV/WAIS-III framework is founded on a multidimensional theory of intelligence that has been supported by clinical research and factor-analytic outcomes. Consequently, both measures provide an assortment of core and supplemental subtests designed to measure selective and distinct features of intelligence (Wechsler, 1997; 2003).

Content. The WISC-IV is comprised of 10 core subtests, from which composite scores can be derived, in addition to five supplemental subtests that extend the range of cognitive abilities sampled; whereas, the WAIS-III encompasses 11 core subtests and three supplemental tests (Wechsler, 1997; 2003). On the WISC-IV, the core subtests are divided across four indices that contribute equally to the Full Scale Intelligence Quotient (FSIQ). The WAIS-III FSIQ, on the other hand, is comprised of only those core subtest that contribute to the verbal and perceptual reasoning indices. The three core subtests that comprise the Verbal Comprehension Index (VCI) on the WISC-IV include Similarities, Vocabulary, and Comprehension. On the WAIS-III, the Comprehension subtest is substituted by the Information subtest. The Perceptual Reasoning Index (PRI) on the WISC-IV is comprised of Block Design, Picture Concepts, and Matrix Reasoning. The

equivalent of the PRI on the WAIS-III, the Perceptual Organization Index (POI), is comprised of Block Design, Picture Completion, and Matrix Reasoning. Digit Span and Letter-Number Sequencing comprise the WISC-IV Working Memory Index (WMI), while the WMI on the WAIS-III incorporates a third, Arithmetic, subtest. Coding and Symbol Search comprise the Processing Speed Index (PSI) on the WISC-IV, while Digit-Symbol Coding and Symbol Search comprise the PSI on the WAIS-III. In addition to subtest and composite scores, process scores providing more detailed information with regard to an individual's performance, are available (Wechsler, 1997; 2003).

The PSI is purported to be a valid measure of information processing speed, which is dynamically related to working memory and higher-order reasoning abilities (Wechsler, 1997; 2003). The WMI is purported to measure an individual's ability to actively maintain information in conscious awareness, manipulate the information, and recall it within a few seconds in a new format. Only selected measures from the PSI and the WMI were utilized in the present study, and are thus subsequently discussed. Specifically, the study incorporated the child and adult versions of Symbol Search and Digit Span. Symbol Search is a core Processing Speed subtest in which the individual is given a specific time frame by which to scan a search group comprised of symbols and indicate whether the target symbol(s) match those in the search group (Wechsler, 1997; 2003). In addition to measuring processing speed, Symbol Search is purported to measure a host of cognitive functions, including short-term visual memory, visual-motor

coordination, cognitive flexibility, attention, and visual discrimination (Flanagan & Kaufman, 2004).

Digit Span is a core Working Memory subtest composed of two parts: Digit Span Forward (DSF) and Digit Span Backward (DSB; Wechsler, 1997; 2003). Digit Span Forward requires the individual to repeat a series of numbers in the same order, after they have been read aloud by the examiner. Digit Span Backward requires that the individual repeat a series of numbers in the reverse order of that presented by the examiner. Each part incorporates several two-trial items of increasing difficulty (Wechsler, 1997; 2003). The Digit Span Forward task involves attention, rote memory, and auditory processing, while Digit Span Backward involves working memory and mental manipulation (Flanagan & Kaufman, 2004). Of note, the WAIS-III does not provide separate normative data for the DSF and DSB; as such, the combined score for Digit Span was utilized in the present study.

Standardization. Normative data for the WISC-IV are based on a sample of 2,200 children between 6 and 17 years of age (Wechsler, 2003). The sample included an equal number of males and females across 11 age groups. A stratified, random sampling design was utilized to ensure the inclusion of a representative proportion of children from each demographic group. U.S. census data from 2000 was utilized for the basis of stratification along the following variables: age, sex, race, parent education level, and geographic region. Additionally, a representative proportion of children (approximately

5.7%) from the special group studies were incorporated into the normative sample to accurately reflect the population of children attending school (Wechsler, 2003).

Normative data for the WAIS-III are based on a sample of 2,450 individuals between the ages of 16 and 89 year (Wechsler, 1997). The sample included an equal number of males and females across the 16 through 64 age bands; the older age groups included more women than men, in proportions consistent with census data. A stratified, random sampling method was used to ensure the inclusion of a representative proportion of adults for each demographic variable. U.S. census data from 1995 was utilized for the basis of stratification along the following variables: age, sex, race/ethnicity, education level, and geographic region. Additionally, information culled from a separate sample of 437 individuals served as oversampling data for future research investigations (Wechsler, 1997).

Reliability. Information regarding the internal consistency of all subtests, process scores, and composite scales on both the WISC-IV and WAIS-III was obtained on the standardization sample using split-half and test-retest procedures (Wechsler, 1997; 2003). Split-half reliability coefficients were obtained on subtests that were of similar length and could be divided in half. The correlation coefficients derived from the split-half procedure was corrected for length using the Spearman-Brown formula (Crocker & Algina, 1986; Li, Rosenthal, & Rubin, 1996). Test-retest reliability coefficients were reported for subtests (i.e., processing speed) in which a timing component precluded the subtest from being split into two parallel forms (Wechsler, 1997; 2003). On the

WISC-IV, average coefficients across age groups ranged from .79 to .90 for core subtests and .79 to .88 for supplemental subtests. Average coefficients on the core and supplemental tests on the WAIS-III ranged from .77 to .93 and .70 to .84, respectively. On both measures, Index coefficients were at or above .90 for all ages, with the exception of Processing Speed coefficients, which ranged from .81 to .90 on the WISC-IV and .86 to .90 on the WAIS III. FSIQ reliability on both measures was excellent, with coefficients at or above .96 for each age group (Wechsler, 1997; 2003).

Test-retest stability for all WISC-IV subtests, process scores, and composite scales was based on a sample of 243 children (n = 18 to 27 per age group; Wechsler, 2003). Time intervals ranged from 13 to 63 days, with an average interval of 32 days. Subtest coefficients, corrected for restriction of range, were generally in the .70s or .80s, whereas Index coefficients ranged from .84 for the WMI (ages 8 to 9) to .95 for the VCI (ages 14 to 16). The FSIQ reliability coefficient was at or above .91 for each age group (Wechsler, 2003). Test-retest stability on the WAIS-III was based on a sample of 394 individuals (n = 30 per age group; Wechsler, 1997). Time intervals ranged from 14 to 84 days, with an average test-retest interval of 34.6 days. Corrected subtest coefficients were generally in the .70s and .80s, with the exception of Vocabulary and Information coefficients, which were noted to be in the .90s. Index coefficients ranged from .83 for the POI (ages 16-29) to .96 for the VCI (ages 55 to 74). The FSIQ was at or above .95 for each age group (Wechsler, 1997).

Validity. Wechsler (1997; 2003) employed a variety of techniques aimed at increasing the content validity of the WISC-IV and WAIS-III. Content revisions, for instance, were conducted to ensure that items and subtests adequately represented the domains of cognitive functioning that the instrument intended to measure. Specifically, results from traditional Mantel-Haenszel bias analysis (Holland & Thayer, 1988) and item response theory (IRT) bias analysis (Hambleton, 1993) were utilized to eliminate potentially problematic items (Wechsler, 1997; 2003). An advisory panel composed of nationally recognized experts in school psychology and clinical neuropsychology was also assembled to work with the research team throughout the development of both measures. Additionally, focus groups and telephone surveys allowed other professionals (e.g., psychometricians, measurement consultants) to rate the research version of the instrument on a variety of aspects affecting validity, such as content and bias. At all stages of test development, modifications were made based on accumulated feedback from experts, as well as quantitative results gathered from pilot and tryout studies (Wechsler, 1997; 2003).

Construct validity across both measures was assessed by examining the intercorrelations among items, subtests, and composite scores (Wechsler, 1997; 2003). In general, subtests contributing to a specific domain or index correlated more highly with each other than with subtests comprising other indexes. Working Memory subtests, for instance, correlated most highly with other WMI subtests (Wechsler, 1997), although results from the WISC-IV indicated moderate correlations between the WMI subtests and

Verbal Comprehension Index subtests (Wechsler, 2003). These findings, however, are expected considering the auditory comprehension demands of the subtests comprising the WMI and VCI. Similarly, Processing Speed subtests correlated most highly with other PSI subtests, although moderate correlations with other subtests (e.g., Block Design) were observed. Overall, these results provide evidence for the sound internal structure of the WISC-IV and the WAIS-III (Wechsler, 1997; 2003).

Evidence of criterion-related validity was ascertained by comparing the WISC-IV and WAIS-III against other scales for children and adults (Wechsler, 1997; 2003). Specifically, the WISC-IV was evaluated against the Wechsler Intelligence Scale for Children-Third Edition (WISC-III; Wechsler, 1991), the Wechsler Preschool and Primary Scale of Intelligence-Third Edition (WPPSI-III; Wechsler, 2002), the WAIS-III (Wechsler, 1997), the Children's Memory Scale (CMS; Cohen, 1997), and the Adaptive Behavior Assessment System-Second Edition (ABAS-II; Harrison & Oakland, 2003). The WAIS-III, on the other hand, was compared against the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981), the Stanford-Binet Intelligence Scale-Fourth Edition (SB-IV; Thorndike, Hagen, & Sattler, 1986), the Standard Progressive Matrices (SPM; Raven, 1976), and the WISC-III (Wechsler, 1991). Overall, comparisons between common composites and subtests resulted in moderate to high correlations (Wechsler, 1997; 2003).

Lastly, in order to compare the diagnostic sensitivity and utility of the WISC-IV and WAIS-III, data was collected on several clinical groups with various developmental,

psychiatric, and neurological conditions, including MS (Wechsler, 1997; 2003). Clinical groups were selected based on their presumed or documented susceptibility to deficits in learning and cognition. Overall, the results of the clinical studies revealed that the WISC-IV and WAIS-III are sensitive to divergent degrees of neurocognitive dysfunction. Strengths and weaknesses within domains were generally consistent with attributes of the respective disorder (Wechsler, 1997; 2003).

California Verbal Learning Test-Children's Version (CVLT-C) and California Verbal Learning Test-Second Edition (CVLT-II)

Overview. The California Verbal Learning Test-Children's Version (CVLT-C; Delis et al., 1994) and the California Verbal Learning Test-Second Edition (CVLT-II; Delis, Kramer, Kaplan, & Ober, 2000) are a pair of clinically-sensitive and comprehensive measures of verbal learning and memory. While the CVLT-C was specifically developed for use with children between the ages of 5 and 16 years, the CVLT-II can be administered to individuals between 16 and 89 years of age. The CVLT-C and CVLT-II are comprised of recall, recognition, and delayed tasks that allow for a brief, individual assessment of the strategies and processes involved in the learning and retrieval of verbal material. Specifically, the CVLT-C provides 27 primary variables that offer information with regard to how verbal learning transpires, or fails to transpire, within the context of an everyday memory task, while the CVLT-II provides standardized scores for over 50 variables. In this way, the CVLT-C and CVLT-II have utility in the

diagnosis and treatment of memory impairments secondary to a host of neurological disorders and psychiatric conditions (Delis et al., 1994; 2000).

Content. The CVLT-C is comprised of two shopping lists (List A and List B) that were strategically constructed to be as comparable as possible (Delis et al., 1994). List A encompasses 15 items that contribute equally to three semantic categories (i.e., clothing, fruits, toys). List B is described as a proactive interference list containing 15 items that contribute equally to the following categories: furniture, fruits, and sweets. Utilizing these lists as the basis for assessment, the CVLT-C is divided into five immediate free-recall trials of List A followed by an interference free-recall trial of List B. Five additional trials utilizing List A are subsequently administered and include: a short-delay free-recall trial, a short-delay cued-recall trial, a long-delay free-recall trial, a long-delay cued-recall trial, and a long-delay recognition trial. The cued response procedure prompts the child by asking for recall of items that are (a) things to wear, (b) things to play with, and (c) fruits. The recognition procedures require the child to indicate whether or not 45 items were contained within the original shopping list (Delis et al., 1994).

The CVLT-II is structured similarly to the CVLT-C with regard to the immediate and delayed free and cued recall trials, with some exceptions. Foremost, the CVLT-II items are not presented in a shopping-list format (Delis et al., 2000). Rather, the examinee is prompted to recall as many items as they can from a list of words. List A is composed of 16 items that contribute equally to four semantic categories (i.e., vegetables, animals, methods of traveling, furniture), while the 16 items on List B contribute equally

to the following categories: vegetables, animals, instruments, and parts of buildings. Like the CVLT-C, the CVLT-II offers a yes/no recognition trial, although an optional forced-choice recognition trial is also available (Delis et al., 2000). For the present study, the following variables from the CVLT-C and CVLT-II were utilized: Trial 1 Correct, Trial 5 Correct, Trials 1-5 Correct, Trial B Correct, Long-Delay Free Recall Correct, and Learning Slope.

Standardization. Normative data for the CVLT-C were derived from a standardization sample of 920 children between the ages of 5 and 16 years who qualified based on specific exclusionary criteria (Delis et al., 1994). The sample included 461 males and 459 females across 12 age groups, with approximately equal representation of both sexes in each age group. A stratified, random sampling design was utilized to ensure the inclusion of a representative proportion of children from each demographic group. U.S. census data from 1988 was utilized for the basis of stratification along the following variables: age, gender, race/ethnicity, parent education level, and geographic region (Delis et al., 1994).

The normative reference sample for the CVLT-II included 1,087 adults between the ages of 16 and 89 years (Delis et al., 2000). More specifically, the sample was comprised of 565 females and 522 males across seven age groups. A stratified, random sampling method was utilized to ensure the inclusion of a representative proportion of children from each demographic group. U.S. census data from 1999 was utilized for the basis of stratification along the following variables: race/ethnicity, education level, and

geographic region. Because gender differences were found to be the second most significant moderating variable after age, normative data for most scores within each of the seven age groups were developed separately (Delis et al., 2000).

Reliability. The reliability of the CVLT-C is based upon estimates of internal consistency and test-retest stability (Delis et al., 1994). Internal consistency was reported across trials, semantic categories, and words. Across-trial consistency was evaluated by utilizing odd-even correlations. This type of internal consistency analysis was utilized given the odd number of trials. The Spearman-Brown formula was subsequently applied to the average of the correlation coefficients culled from the odd-even procedure. Odd-even reliability coefficients ranged from .84 to .91, with an averaged coefficient, based on Fischer's *z* transformation, of .88. Coefficient alpha internal consistency reliability coefficients were comparable, ranging from .81 to .88, with an averaged coefficient of .85. Across semantic category consistency was determined by evaluating performance across item sets that were semantically unrelated. This method involved using the List A recall scores across the five learning trials and dividing the words into three semantic categories (i.e., fruits, things to wear, and things to play with). Each semantic group was comprised of 10 stimulus items with five words in each category (e.g., fruits and things to wear, fruits and things to play with, things to wear and things to play with; Delis et al., 1994).

Alpha coefficients were calculated based upon the three category scores. Reliability coefficients ranged from .64 to .80, with an average of .72. Across word

consistency was based on the total (five-trial) score for each of the 15 stimulus words on List A. The average odd-even correlation was .83, and the average coefficient alpha correlation was .81 (Delis et al., 1994). Cronbach's alpha coefficients of .70 or above typically indicate acceptable reliability (Lodico, Spaulding, & Voegtle, 2006). Given this criterion, the CVLT-C Trials 1-5 seem to possess a high degree of internal consistency that support the acquisition of reliable estimates of recall ability.

To determine the stability of performance on the CVLT-C, a sample of 106 children between the ages of 8 and 16 years completed the CVLT-C on two occasions, with an interval range of 10 to 42 days, and a median test-retest interval of 28 days (Delis et al., 1994). Some improvement in recall of word-list items from both lists was noted for all ages. Comparable increases in perseveration were noted across all age groups. In contrast, only 12-year-old children demonstrated improvement on recognition hits, and intrusion rates remained unchanged (Delis et al., 1994).

The reliability of the CVLT-II was reported using three varying forms of split-half reliability (Delis et al., 2000). The first method evaluated across-trial consistency by utilizing odd-even correlations (Trial 1 + 3 versus Trial 2 + 4, and Trial 2 + 4 versus Trial 3 + 5). The Spearman-Brown formula was subsequently applied to the average of the correlation coefficients. The reliability estimates within age groups were largely above .90. Reliability for the total sample was excellent (r = .94), and no significant differences between males and females were noted. This same method was also applied to a mixed clinical sample of 124 neuropsychiatric patients. The split-half reliability estimate for this

sample was equally robust (r = .96). The second approach to estimating internal consistency involved the examination of participants' performance in the four categories of words in List A across the five immediate-recall trials. The total items recalled across the five trials on each of the four categories were treated as a variable. These variables were then treated as items and coefficient alphas were calculated. The reliability coefficients for the standardization and mixed clinical samples were .82 and .83, respectively. In the last method of calculating internal consistency, the number of times each of the 16 words from List A were recalled across the five immediate trials was examined. Split-half reliability coefficients for the standardization and mixed clinical samples were .79 and .83, respectively (Delis et al., 2000).

To determine the stability of scores on the CVLT-II, the measure was administered to a sample of 78 individuals on two occasions, with an interval range of 9 to 49 days, and a median test-retest interval of 21 days (Delis et al., 2000). On average, participants recalled approximately eight more words across the five learning trials upon retest. Lastly, using a counterbalanced method, the CVLT-II Standard Form and the CVLT-II Alternate Form were administered to a sample of 288 nonclinical individuals. All of the correlation coefficients were found to be statistically significant at $p < .001$, with reliability coefficients ranging between .72 and .79 (Delis et al., 2000).

Validity. Three major lines of evidence for the validity of the CVLT-C and CVLT-II are available and include content, construct, and criterion validity (Delis et al., 1994; 2000). The theoretical foundation of the CVLT-C and CVLT-II and the extent to

which they adequately capture the various aspects of verbal learning and memory is based upon a rich and solid body of literature that has evolved through several decades. Based on the elements of scientific research and brain-behavior relationships, this body of literature has described the various strategies and patterns associated with learning in addition to the range of problems related to encoding and retrieval, both immediately and across time. Given the intricacies and complexities of learning and memory, Delis et al. (1994; 2000) employed this body of literature in the design of the CVLT-C and CVLT-II in order to create multifaceted measures that would assess the various component skills that have been found to carry the most clinical relevance. Additionally, the validity of the CVLT-II is largely grounded in its relationship to the original California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Ober, 1987), which has accumulated a significant amount of research support (Delis et al., 2000). The construct validity of the CVLT has, in fact, been demonstrated in over 200 publications with both non-clinical samples and cohorts with a wide range of neurological, psychiatric, and medical conditions. The high degree of concurrent validity between the CVLT-II and the CVLT (r = .72 to .86) suggests that the CVLT-II maintains a similar level of construct validity as its predecessor (Delis et al., 2000).

To confirm the internal structure of the CVLT-C and CVLT-II and thus ensure construct validity, Delis et al. (1994; 2000) performed independent exploratory factor analyses on the 19 indices of both measures. Utilizing a principal components technique, factors with eigenvalues greater than 1 were retained, and loadings greater than .40 were

considered significant. On both measures, the varimax rotated factor structure of the 19 indices yielded a six-factor solution, which is comparable to the underlying structure of the original CVLT. In general, these results indicate that the indices of the CVLT-C and CVLT-II cluster into theoretically meaningful factors that are consistent with the constructs they were designed to measure (Delis et al., 1994; 2000).

Given the significant, though modest, correlation between verbal intelligence and verbal memory, Delis et al. (1994; 2000) compared performance on the CVLT-C and CVLT-II against performance on the Vocabulary subtest of the Wechsler Intelligence Scale for Children-Revised (WISC-R; Wechsler, 1974) and the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), respectively. Correlations between the CVLT-C learning trial scores and the WISC-R (Wechsler, 1974) Vocabulary standard scores were noted to range from .32 to .40, indicating a significant, yet mild relationship. Correlations between the CVLT-II learning trial scores and the WASI (Wechsler, 1999) Vocabulary standard scores was equivalent (r = .43). Given the low percentage of shared variance, findings support the utility of the CVLT-C and CVLT-II in evaluating a domain of cognition that is different from verbal intelligence (Delis et al., 1994; 2000).

Delis-Kaplan Executive Function System (D-KEFS)

Overview. The Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2001) is a compilation of nine stand-alone measures of verbal and nonverbal executive functions in children and adults between the ages of 8 and 89 years. The D-KEFS represents the first grouping of executive functioning tests co-normed on a

large stratified sample (Homack, Lee, & Riccio, 2005). The D-KEFS is largely based on a variety of procedures that have been empirically demonstrated to be efficacious in the detection of executive dysfunction (Delis et al., 2001). In this way, the D-KEFS is empirically-based rather than theoretically-derived. Although some of the tests within the D-KEFS are unique, many are modified from notorious and widely supported tests of executive functioning (e.g., Stroop test, Tower task). In order to avoid extensive and lengthy evaluations, clinicians are able to choose the most appropriate tests to administer, based on individual needs (Baron, 2004). Moreover, because several tasks involve multiple conditions and levels, clinicians are able to obtain a host of process scores in addition to primary scores (Delis et al., 2001).

Content. The D-KEFS is a compilation of nine stand-alone tests that were individually designed to assess a wide spectrum of verbal and nonverbal executive functions (Delis et al., 2001). Eight of these tests are standardized for administration to children and include the following: Word Context Test, Sorting Test, Twenty Questions Test, Tower Test, Color-Word Interference Test, Verbal Fluency Test, Design Fluency Test, and Trail Making Test. For the current study, only the Verbal Fluency Test was utilized, and is thus discussed in more detail. The Verbal Fluency Test is a modified version of the Controlled Oral Word Association Test (COWAT; Benton and Hamsher, 1976; Benton, Hamsher, & Sivan, 1994) and consists of three conditions: Letter Fluency, Category Fluency, and Category Switching. In Letter Fluency (Condition1), the examinee is asked to state words that begin with a particular letter as quickly as possible. In

Category Fluency (Condition 2), the examinee is asked to verbalize words of a specific semantic category as quickly as possible. In Category Switching (Condition 3), the examinee is required to alternate saying words of two different categories as quickly as possible. The Verbal Fluency Test is purported to measure fluent productivity in the verbal domain and cognitive flexibility (Delis et al., 2001). For the current study, only the letter fluency (FAS) task, which measures "speeded lexical production and the degree of automatic lexical access" (Miller, 2010, p. 254), was utilized. Specifically, the study analyzed the Total Correct score.

Standardization. The D-KEFS was standardized on a nationally representative sample of 1,750 individuals between the ages of 8 and 89 years (Delis et al., 2001). Stratification was based on the 2000 U.S. census for age, sex, race/ethnicity, years of education, and geographic region. The sample included an equal number of males and females across 16 age groups, ranging from 8 to 89 years of age. Examinees were excluded from the study if they endorsed any psychiatric or medical condition that could potentially affect performance on cognitive tests. Examiners were selected based upon experience with psychometric testing, certification, and licensure (Delis et al., 2001).

Reliability. Internal consistency for the majority of the D-KEFS tests, including the Verbal Fluency Test, was assessed using split-half and test-retest reliability procedures. Overall, the split-half reliability estimates were widely varied across tests and age groups (Delis et al., 2001). Split-half reliabilities specific to the Verbal Fluency Test were in the low to high ranges (r = .43 to .90), although moderate to high reliability

coefficients (r = .68 to .90) were reported for Letter Fluency. Test-retest reliability coefficients were based on a sample of 101 examinees from all of the age groups. Time intervals ranged from 9 to 74 days, with an average interval of 25 days. For many of the tests, practice effects were often observed, which lead to low test-retest reliabilities. The test-retest measures for the Verbal Fluency test, however, were noted to have moderate to high reliability. Specifically, test-retest reliability coefficients for Letter Fluency ranged from .67 to .88 (Delis et al., 2001).

Validity. Evidence of validity for the D-KEFS was reported from various sources, including intercorrelations between measures provided by each of the D-KEFS tests, correlations between the D-KEFS tests and established neurocognitive instruments, and data from studies assessing the sensitivity of the D-KEFS tests to numerous clinical populations (Delis et al., 2001). Overall, intercorrelations for the D-KEFS tests were indicative of adequate convergent and discriminate validity, as evidenced by expected positive and negative correlations (Homack et al., 2005). For instance, while Letter Fluency, Category Fluency, and Category Switching were moderately associated with each other, Category Switching and Category Fluency were reported to have a stronger association (Delis et al., 2001).

Information regarding validity was also gathered from correlational studies between the D-KEFS, the CVLT-II (Delis et al., 2000), and the Wisconsin Card Sorting Test (WCST; Heaton, 1981). Convergent validity was evidenced between the D-KEFS and the WCST scores, which were reported to have moderate correlations. These results

generally indicate that the D-KEFS and WCST share some degree of variance in measuring higher-order executive functions, while also contributing unique variance in the assessment of different aspects of executive functioning. Discriminate validity, on the other hand, was evidenced between the D-KEFS and the CVLT-II. Specifically, the CVLT-II recall and recognition measures had low positive correlations with correct responses across all conditions of the Verbal Fluency Test. Finally, investigations utilizing the D-KEFS have supported its ability to accurately differentiate the neuropsychological profiles of individuals with Alzheimer's disease and Huntington's disease (Delis et al., 2001). Similar findings have been reported on individuals diagnosed with fetal alcohol exposure, schizophrenia, and chronic alcoholism (Homack et al., 2005).

Beery-Buktenica Developmental Test of Visual-Motor Integration, 5th Edition (Beery VMI)

Overview. The Beery-Buktenica Developmental Test of Visual-Motor Integration, 5th Edition (Beery VMI; Beery & Beery, 2006) and its two supplemental counterparts, the Visual Perception Test and Motor Coordination Test, provide a comprehensive, yet efficient screening of visual-motor abilities. The Beery VMI encompasses a developmental sequence of geometric forms designed to assess the extent to which individuals can integrate their visual and motor abilities. Although the Beery VMI was originally developed as a screener for children between 2 and 18 years of age, it has become a valuable assessment tool for adults up to 100 years of age. The Beery VMI

has been utilized across countries and is purported to be virtually culture-free (Beery & Beery, 2006).

Content. The Beery VMI encompasses a sequence of geometric figures that are to be imitated or copied utilizing paper and pencil (Beery & Beery, 2006). The fifth edition of the Beery VMI offers three forms: the Full Form, the Short Form, and the Adult Form, all of which can be administered individually or in a group format. The Full Form can be utilized with children and adults across all age ranges. Although the administration instructions were written for children, no modifications are needed for administration with adults. The Full Form is comprised of 30 items, including three imitation/copy items and three spontaneous drawing/scribbling items. The Short Form, on the other hand, is intended for use with children between 2 and 7 years of age. The Short Form is equivalent to the Full Form, except that it contains only the first 15 Beery VMI forms, for a total of 21 items. The Adult Form was designed for adults between the ages of 19 and 100 years of age. The Adult Form is similar to the Full Form except it omits items 1 through 6, which are intended for children who are under 5 years of age. Two standardized supplemental subtests are also available and include the Beery VMI Visual Perception test and the Beery VMI Motor Coordination test. These tests use the same geometric forms as the Beery VMI and are intended to provide additional information with regard to visual and motor contributions to Beery VMI performance (Beery & Beery, 2006). In the present study, the Beery VMI Full Form and the supplemental Visual Perception test were utilized.

Standardization. Normative data for the Beery VMI were originally based on a sample of 1,030 children from Illinois (Beery & Beery, 2006). Since 1964, the Beery VMI has been standardized four times with more than 9,000 children. Specifically, the Beery VMI was cross-validated in 1981 with 2,060 children in California, in 1989 with a nationally representative sample of 2,743 children, and in 1995 with 2,614 children from the five major U.S. regions. The fifth edition of the Beery VMI is largely based on information culled from a norm study conducted in 2003 on 2,512 children between 1 and 18 years of age. In order to adequately represent 2000 U.S. census data, samples for the 2003 study were selected from across 23 childcare, preschool, public, and private entities. The following demographic variables were closely approximated: gender, age, ethnicity, residence (i.e., urban, non-metropolitan), geographic region, and parent education. Additionally, the Beery VMI was nationally standardized in 2006 on a sample of 1,021 adults between the ages of 19 and 100 years. Participants were culled from 27 states, representing all major geographic regions within the U.S. The norming sample closely approximated 2000 U.S. census data along the following variables: gender, age, ethnicity, residence, geographic region, and education. Of note, the Beery VMI norms have remained consistent over time and place, particularly as it relates to preschool and elementary grade levels (Beery & Beery, 2006).

Reliability. The reliability of the Beery VMI is based upon estimates of internal consistency, test-retest stability, and interrater reliability. Internal consistency was evaluated by utilizing odd-even correlations and coefficient alpha internal consistency

estimates (Beery & Beery, 2006). For children and adolescents between 2 and 17 years of age, odd-even split half correlations ranged from .82 to .90, with an average coefficient of .88; whereas, coefficient alpha reliability ranged from .79 to .89, with an average coefficient of .82. While odd-even split half correlations for adult samples were not reported, coefficient alpha internal consistently estimates were noted to range from .85 to .94, with an average coefficient of .89 (Beery & Beery, 2006).

To determine the stability of performance on the Beery VMI, a sample of 115 children between the ages of 5 and 11 years and a sample of 20 adults between the ages of 60 and 69 years completed the Beery VMI on two occasions (Beery & Beery, 2006). Time between administrations averaged 10 days for the child sample and 5 days for the adult sample. The overall test-retest coefficients for the child and adult samples were .89 and .88, respectively. Moreover, in order to ascertain interscorer reliability, two individuals independently scored 100 randomly-selected Beery VMI child protocols and 25 adult protocols from the norming group. Interscorer reliability coefficients of .92 and .94 were reported for the child and adult samples, respectively. According to test authors, the average of Anastasi's (1988) three major reliability error sources (i.e., interscorer, internal consistency, and test-retest) provides the best indication of overall reliability. Based on this criteria, the Beery VMI possesses an overall reliability of .92 (Beery & Beery, 2006).

Validity. The content validity of the Beery VMI is grounded in the extensive empirical research, item-analyses, and revisions conducted by its test authors (Beery &

Beery, 2006). Evidence of successful item selection based on criteria that could affect validity are evidenced by the Rasch-Wright (Wright & Stone, 1979) item analysis results. Rasch-Wright analysis was utilized to assess the fit between test items and test constructs, the unique contribution of each test item to the test, and the items' ability to differentiate among individuals (Beery & Beery, 2006). Overall, results indicated high content validity. In order to provide evidence for the construct validity of the measure, test authors analyzed the intercorrelations between the Beery VMI and its supplemental tests. Although average correlations were relatively inadequate (r =.28 to .41), they were significant beyond the .05 level of confidence. Information regarding the concurrent validity of the fifth edition of the Beery VMI was not reported. Rather, concurrent validity was based on studies comparing the fourth edition of the Beery VMI against the following measures: the Copying subtest from the Developmental Test of Visual Perception (DTVP-2; Hammill, Pearson, & Voress, 1993), the Drawing subtest from the Wide Range Assessment of Visual Motor Abilities (WRAVMA; Adams & Sheslow, 1995), and the Bender-Gestalt (Bender, 1938). Correlations between the fourth edition of the Beery VMI and these measures were adequate (r = .75, .52, and .56, respectively; Beery & Beery, 2006).

Grooved Pegboard Test (GPT)

Overview. The Grooved Pegboard Test (GPT; Klove, 1963) is a brief and widely used measure of manual dexterity and hand-eye coordination (Baron, 2004). The GPT has been an integral component of various batteries, including the Wisconsin Motor

Steadiness Battery (Matthews & Klove, 1964), the extended Halstead Reitan Neuropsychological Battery (HRNB; Halstead, 1947; Reitan, 1969; Reitan and Davison, 1974), and the Repeatable Cognitive-Perceptual Motor Battery (Lewis & Kupke, 1992). While the GPT is primarily believed to measure manual dexterity, psychomotor speed, and fine motor coordination, it has been utilized to make inferences with regard to brain dysfunction. More specifically, performance scores have been purported to reflect degrees of lateralized cerebral dysfunction (Baron, 2004; Mitrushina, Boone, Razan, & D' Elia, 2005).

Content. The GPT consists of 25 ridged pegs and a 10-centimeter metal board containing a five-by-five matrix of randomly positioned holes (Baron, 2004). Using one hand at a time (dominant precedes nondominant), the examinee is instructed to lift each keyhole-shaped peg from the storage well and insert it into one of the randomly oriented holes as quickly as they can. The holes must be filled in a row-by-row fashion and only one peg can be inserted into the board at a time. Children between the ages of 5 and 9 years are required to fill only the two upper rows with each hand; whereas, individuals who are 9 years and above complete an entire 25-hole trial with each hand. Timing begins when the individual commences the task, and each trial can be discontinued after 5 minutes. The GPT provides three scores: time required to complete each trial, the number of pegs dropped in each trial, and the number of pegs correctly inserted in each trial (Baron, 2004). For the present study, the score associated with completion time for both the dominant and nondominant hands was utilized.

Standardization. Information regarding the normative data of the GPT is largely based on various independent investigations. Knights and Moule (1968), for instance, published normative data compiled from a sample of 184 children between the ages of 5 and 14 years. The sample included 96 students from a Class II elementary school and 88 students from a Class IV elementary school; class types were based on the Two-Factor Index of Social Position (Hollingshead, 1957). While information with regard to the demographic characteristics of the sample were not reported, IQ scores for the Class II and Class IV participants were noted to range from 95 to 133 and 80 to 126, respectively (Knights & Moule, 1968). These data were subsequently updated to reflect "smoothed" data. Specifically, since the raw data plotted across age bands revealed sawtooth patterns, the means and standard deviations were statistically smoothed to provide consistent standard scores across age levels (Baron, 2004). Another set of normative data derived from Trites (1977), provides the means and standard deviations associated with test completion time for males and females between 5 and 60+ years of age; however, no information with regard to sample size or sample characteristics was reported.

More comprehensive adult norms have been derived from a sample of 365 Canadian residents between 18 and 69 years of age (Bornstein, 1985). In an effort to obtain a representative cross-section of the community, recruitment notices were posted across university campuses, unemployment offices, and local newspapers. Moreover, the sample was stratified by sex, age, and level of education (Bornstein, 1985). A more recent study (Heaton, Grant, & Matthews, 1991) provides normative data based on 486

adults from across the U.S. and Canada. Data were collected over a 15-year period through multicenter collaborative efforts. The sample was comprised of 316 males and 170 females with a mean education level of 13.6 years and a mean FSIQ of 113.8. For the present study, norms derived from Knights and Moule (1968) were utilized for patients between the ages of 5 and 14 years, while norms derived from Trites (1977) were used for those between 15 and 19 years of age.

Reliability. While information regarding the reliability of the GPT is limited, Knights and Moule (1968) have provided some evidence of test-retest reliability. Data was gathered from a sample of 40 children between the ages of 8 and 15 years who were suspected of having a neurological dysfunction and thus enrolled in a 6-week methylphenidate trial. All participants completed the GPT before commencing the trial and after completing the trial. Reliability coefficients for completion time were .80 and .81 for dominant and nondominant hands, respectively. Reliability estimates for number of errors were much lower (dominant hand: $r = 20$; nondominant hand: $r = .21$; Knights & Moule, 1968). A separate study reported test-retest reliability coefficients ranging from .69 to.76 for the dominant hand and .68 to .78 for the non-dominant hand over a six-month period (Ruff & Parker, 1993). While the information culled from these studies is valuable, investigations addressing the reliability of the GPT with demographically diverse samples are warranted.

Validity. Although there is a dearth of information regarding the content and construct validity of the GPT, several studies have documented its diagnostic sensitivity

and utility. In a study conducted by Knights et al. (1991), for instance, the GPT was shown to successfully discriminate severely impaired children from those with mild and moderate impairments. Specifically, a sample of 76 children were divided into three groups on the basis of head-injury severity defined by the Glasgow Coma Scale (Teasdale & Jennett, 1974) in addition to duration of intracranial pressure. Participants were successively assessed at three discrete time intervals: at time of hospital discharge, 3 month post-discharge, and 9 months post-discharge. While several differences in cognitive functioning related to severity of injury were observed, they were most notable in the areas of performance IQ and visual-motor speed/dexterity (Knights et al., 1991). Moreover, in a study conducted by Matthews, Cleeland, and Hopper (1970), 30 patients with MS and 30 patients with other CNS disorders were assessed utilizing a battery comprised of various cognitive, sensory-motor and personality measures. Cohorts were matched on the basis of sex, age, gender, education level, and FSIQ. While the two cohorts exhibited comparable performance across measures of concept formation, verbal intelligence, and speech discrimination, performance on the GPT was significantly lower for MS patients than "control" patients. Specifically, MS patients had a mean completion time of 323.4 seconds; whereas, control patients had a mean of 117.77 seconds (Matthews et al., 1970).

Trail Making Test (TMT)

Overview. The Trail Making Test (TMT) was originally constructed in 1938 and included in the Leiter-Partington Adult Performance Scale (Partington & Leiter, 1949) as

"Partington's Pathways." The TMT was a component of the Army Individual Test Battery (Army Individual Test Battery, 1944; Armitage, 1946) and subsequently added to the HRNB (Halstead, 1947; Reitan, 1958; Reitan, & Davison, 1974). In 1971, the TMT was adapted for use with children (Reitan, 1971). Given its sensitivity to general brain dysfunction and cognitive impairment across domains (Lezak, 1995; Reitan, 1971), the TMT has become one of the most widely used neuropsychological tests. The TMT has been described as a multifaceted measure with factor analytic loadings across various factors including visual perception, visuomotor speed, visual scanning, sustained attention, and problem-solving skills (Baron, 2004).

Content. The TMT is a timed paper-and-pencil test that encompasses two parts (Baron, 2004; Mitrushina et al., 2005). Part A is comprised of sequentially numbered circles randomly scattered across a page. In order to complete Part A, the examinee must draw a line connecting the numbers in order as quickly as possible. Part B requires the examinee to draw a line while sequentially alternating between numbers and letters as quickly as possible. On both parts, the examinee is immediately prompted to correct any error. In general, Part A and Part B differ in terms of length of trail and perceptual complexity. Moreover, the sequence alternation demands of Part B tap additional executive processes (Baron, 2004; Mitrushina et al., 2005).

The adult version of the TMT is comprised of 25 numbers on Part A and a total of 25 numbers and letters on Part B (Baron, 2004; Mitrushina et al., 2005). This version is appropriate for individual who are 15 years and older. The child or "intermediate"

version, on the other hand, is comprised of 15 numbers on Part A and a total of 15 numbers and letters on Part B. This version is intended for use with children between the ages of 9 and 14 years. The TMT provides two scores, each reflecting the total time of completion for each part. The number of errors committed by the examinee can be utilized for qualitative purposes, while a B/A ratio score can be utilized to make inferences with regard to cognitive efficiency (Baron, 2004; Mitrushina et al., 2005). For purposes of this study, only scores associated with time to completion for Part A and Part B were analyzed.

Standardization. Normative data for the TMT have been derived from various independent studies. Part A and Part B completion time data using the HRNTB, for instance, were presented in an early study by Spreen and Gaddes (1969). The sample was stratified by age and gender and included 272 neurotypically developing children between the ages of 8 and 13 years. Normative data for children between the ages of 9 and 14 years was gathered by Knights (1966). Although the study and normative data are cited in neuropsychological compendiums, information regarding the sample and stratification procedures were not accessible. More recently, Findeis and Weight (1994) culled normative data from a review of 20 articles published between 1965 and 1990. Studies with limited statistical information and scoring ambiguities were excluded. The final selection included 33 measures from both tests. The sample of 3, 225 children between the ages of 5 and 14 years included 910 males, 805 females, and 1510 children of unspecified gender. Participants were primarily Caucasian and of middle to

upper-middle socioeconomic status. Within the context of this study, data specific to the TMT were based on 417 children between 9 and 13 years of age (Findeis & Weight, 1994).

Adolescent and adult normative data were culled from a sample of 193 Canadian participants between the ages of 15 and 64 years (Fromm-Auch & Yeudall, 1983). The sample was comprised of 82 females and 111 males who were recruited through posted advertisements. Mean years of education and IQ were 14.8 and 119.1, respectively (Fromm-Auch & Yeudall, 1983). Additional normative data were gathered from a sample of 90 participants, aged 16 to 69 years (Stuss, Stethem, & Pelchat, 1988). Participants were collapsed into three age groups (16 to 29, 30 to 49, and 50 to 69), with equal gender distribution across groups. Mean years of education for the three groups were 14.1, 14.9, and 13.2, respectively (Stuss et al., 1988). A more recent investigation by Tombaugh (2004) provides normative data for 911 (408 males, 503 females) community-dwelling adults between 18 and 89 years of age. The sample was gathered using various recruitment techniques and stratified by age, gender, and level of education (Tombaugh, 2004). For the present study, norms derived from Knights (1966) were utilized for patients between the ages of 9 and 14 years, while norms derived from Fromm-Auch and Yeudall (1983) were used for those between 15 and 19 years of age.

Reliability. The internal consistency of the TMT has been addressed across numerous studies using test-retest procedures. Earlier studies by Dye (1979), and Stuss, Stethem, and Poirier (1987) demonstrated significant practice effects resulting from short

testing intervals (i.e., one-week), while Lezak (1982) found significant practice effects for Part B over a 6-month interval. In this latter study, test-retest reliability was reported as .98 for Part A and .67 for Part B. Snow, Tierney, Zorzitto, Fisher, and Reid (1988) reported a 12 month test-retest reliability of .64 for Part A and .72 for Part B in 100 adults with a mean age of 67 years. Goldstein and Watson (1989) reported similar coefficients (r = .69 to .94 for Part A; r = .66 to .86 for Part B) in a mixed clinical sample of adults with various neurological conditions. While these studies provide valuable information, investigations addressing the internal consistently of the TMT with pediatric populations is warranted.

Estimates of interrater reliability have also been reported. In 1982, Lezak (1982) hypothesized that the administration and scoring procedures for the TMT resulted in diminished reliability. On this premise, Fals-Stewart (1992) recruited 39 undergraduate students, who were instructed to make specific errors on Part A and Part B of the test. The four psychometricians who participated in the study were gathered from various testing sites and had between 5 and 7 years of experience administering and scoring the TMT as part of the HRNB. The psychometricians tested each participants over the course of two days. Interrater reliability coefficients on both parts was high (r = .94; r =.90), indicating that scorer variance is partially attributable to years of testing experience. In a more recent study examining the cognitive and emotional effects of bilateral anterior cingulate cortex (ACC) lesions, for instance, concordance coefficients of .98 for Part A and .67 for Part B were reported (Cohen et al., 2001).

Validity. The validity of the TMT is primarily grounded in its clinical sensitivity and utility, which has been demonstrated across numerous studies. The TMT has been reported to be sensitive to various clinical populations, including adults with traumatic brain injuries (des Rosiers & Kavanagh, 1987), substance abuse (McCaffrey, Krahula, Heimberg, Keller, & Purcell, 1988), human immunodeficiency virus (Di Sclafani et al., 1997), and Alzheimer's disease (Lafleche & Albert, 1995). A large 1-year follow-up study, for instance, found significant differences in performance on Part A and Part B in a sample of 436 who had sustained head injuries, when compared to 132 controls with injuries to other body parts (Dikmen, Machamer, Winn, & Temkin, 1995). Crockett, Hurwitz, and Vernon-Wilkinson (1990) also found significant differences between patients with anterior and posterior brain dysfunction and a mixed psychiatric sample; these results, however, were primarily related to Part A of the TMT.

Information with regard to the criterion-related validity of the TMT has been gathered from studies comparing the TMT against the Color Trails Test (CTT). Correlations between Trails 1 and Trails 2 of the CTT and Part A and Part B of the TMT have been reported as .41 and .50, respectively (Maj, D'Elia, & Satz, 1993). In a study involving pediatric patients between the ages of 5 and 16 years, the respective parts of the CTT and TMT were noted to correlate at the higher rates of .74 and .69 (Williams et al., 1995). Another study (Ehrenstein, Heister, & Cohen, 1982) examined the divergent validity of the TMT by comparing it against verbal tests with scanning components, such as the Peabody Picture Vocabulary Test (PPVT; Dunn, 1959).

Studies addressing the construct validity of the TMT are not as prevalent, although various factor analytic studies (e.g., Groff & Hubble, 1981; Shum, McFarland, & Bain, 1990) demonstrate that both parts of the TMT load onto similar factors, including visual scanning and perception, visuomotor speed, sustained attention, and problem-solving. The results suggest that Part A and Part B measure similar basic cognitive functions. Other studies, however, (e.g., Heilbronner, Henry, Buck, Adams, & Fogle, 1991) have found that Part A and Part B correlate at .49 with each other, suggesting that each part also measures somewhat different functions. Specifically, Part B is purported to tap additional executive processes, such as cognitive shifting (Baron, 2004; Mitrushina et al., 2005).

Symbol Digit Modalities Test (SDMT)

Overview. The Symbol Digit Modalities Test (SDMT; Smith, 1973; 1982) is a measure of visual scanning, visual tracking, and sustained visual attention that requires efficient and accurate information processing. The SDMT is intended for use with children and adults between 8 and 78 years of age. Performance scores on the SDMT reflect the integration of various neuropsychological processes, including visual, motor, and mental functions. As such, the SDMT is a brief method by which to screen for learning disorders and other cerebral dysfunctions. The SDMT can also be used as a diagnostic tool when used in conjunction with other standardized neuropsychological measures. Because the SDMT is comprised of geometric figures and numbers, it has

limited cultural bias and can thus be administered to individuals who speak a language other than English (Smith, 1982).

Content. The SDMT is a simple, 90-second substitution task that involves the pairing of numbers and geometric symbols (Smith, 1982). The measure offers a written version in which the examinee is required to record their responses on an answer sheet and an oral version in which the examinee is required to verbally state their responses. On both versions, the examinee is given a key containing the numbers 1 through 9, each of which is paired with a distinct geometric symbol. On the written version, the examinee refers to the key in order to rapidly record the number that is paired with each given symbol. On the oral version, the examinee is instructed to use the key in order to state the correct number response; in the latter version, the examiner records the examinee's responses on the answer sheet (Smith, 1982). For purposes of this study, the oral version was utilized.

Standardization. Pediatric normative data for the SDMT was derived from a sample of 3,680 children between the ages of 8 and 17 years (Smith, 1982). The sample included 1,874 males and 1,806 females who were enrolled in regular education classes across 15 public schools in the Omaha area. Schools were selected from low, middle, and upper income neighborhoods to ensure the inclusion of a representative proportion of children from each socioeconomic group. Written scores on the SDMT were obtained for 2,101 of the children in the sample; whereas, oral scores were derived for all children in the study. The written portion was administered using a group format, while the oral

portion was individually administered. Additionally, a sample of 1,579 children were given only the oral portion of the SDMT (Smith, 1982).

Adult norms for the SDMT were obtained from two separate samples (Smith, 1982). The first sample, provided by Centofanti (1975) included 420 adult volunteers (206 males and 214 females) from Florence Township, New Jersey and Madison Heights, Michigan. Individuals with reported neurological involvement and apparent handicaps were excluded from the study. The second sample of 887 volunteers was collectively culled by a group of psychologists in Ann Arbor, Michigan. Participants from both samples were equally divided into six age groups: 18-24, 25-34, 35-44, 45-54, 55-64, and 65 and older. Each age group was subsequently divided into a high school level (12 years of education or less) and a college level (13 years of education or more; Smith, 1982).

Reliability. The reliability of the SDMT is based upon adult studies of test-retest stability and equivalent forms consistency (Smith, 1982). Estimates of test-retest stability were obtained from a study on 80 adults who completed two administrations of the oral and written versions of the test with an average test-retest interval of 29.40 days. Results revealed a .80 correlation for the written version and a .76 correlation for the oral version, indicating adequate stability of SDMT scores over time. This same study was utilized to examine equivalent forms consistency. Correlation estimates of .82 and .84 were obtained between the written and oral forms at initial testing and retesting. A similar study on 887 adults revealed a correlation of .78. Combined, the equivalent forms studies provide evidence of the compatibility of the two versions (Smith, 1982).

Validity. In order to validate the measure's normative data, SDMT norms have been compared against SDMT scores derived from nonclinical groups. A study by Krull (1968), for instance, compared the SDMT norms with scores obtained from 147 (74 males and 73 females) public school students between the ages of 9 and 19 years of age. Overall, mean scores were comparable to those obtained from the standardized sample (Krull, 1968). The SDMT has also been administered to an array of diverse clinical populations in order to ascertain its diagnostic sensitivity and utility. Investigations of this nature have supported the SDMT's ability to accurately identify children and adults with various types of neurological conditions, including traumatic brain injury, aphasia, mild intellectual disability, and Huntington's disease. Results of these studies have routinely demonstrated that clinical groups produce significantly subnormal scores on the written and oral versions of the SDMT when compared to nonclinical groups (Smith, 1973, 1982).

Behavior Assessment System for Children-Second Edition (BASC-2)

Overview. Behavior Assessment System for Children-Second Edition (BASC-2; Reynolds & Kamphaus, 2004) is a comprehensive, multidimensional assessment tool that utilizes rating scales to evaluate the emotional and behavioral functioning of children and young adults between the ages of 2 and 21. The BASC-2 was designed to facilitate the differential diagnosis and educational classification of various disorders through a triangulated method that involves the examination of behavior in multiple settings (e.g., home, school), the evaluation of the child's self-perceptions, and the inclusion of relevant

background information. The BASC-2 is highly relevant to specific categories of disorders as outlined in the Diagnostic and Statistical Manual of Mental Disorders-Text Revision (DSM-IV-TR; American Psychiatric Association, 2000) as well as those addressed by the Individuals with Disabilities Education Act (IDEA, 1997; Reynolds & Kamphaus, 2004).

Content. The BASC-2 offers a Teacher Rating Scale (TRS), a Parent Rating Scale (PRS), and a Self-Report of Personality (SRP; Reynolds & Kamphaus, 2004). The TRS measures both adaptive and problem behaviors in the school setting. The TRS forms were designed to target three age levels: preschool (2 through 5 years), child (6 through 11 years), and adolescent (12 through 21 years). The form contains behaviors that the respondent rates on a four-point Likert scale based on frequency (*Never* to *Almost always*). The preschool form contains 100 items; whereas, the child and adolescent forms contain 139 items. These items assess the broad domains of Externalizing Problems, Internalizing Problems, School Problems, and Adaptive Skills. The PRS measures adaptive and problem behaviors in the home and community settings. Like the TRS, the PRS uses the same four-choice response format to target three age levels. However, the PRS is comprised of 134 preschool items, 160 child items, and 150 adolescent items. Moreover, the PRS includes an Activities of Daily Living scale that the TRS does not measure, although it does not offer a School Problems composite (Reynolds & Kamphaus, 2004).

The SRP is an inventory of statements that respondents answer in one of two ways; some items require a bimodal (*True/False*) response, while others require a frequency-based response (Never to Almost always; Reynolds & Kamphaus, 2004). The SRP targets three age levels: child (8 to 11 years), adolescent (12 to 21 years), and young adult (18 to 25 years). The child form contains 139 items; whereas, the adolescent and young adult forms contain 176 and 185 items, respectively. In children and adolescents, these items produce a School Problems, Internalizing Problems, Inattention and/or Hyperactivity, Personal Adjustment, and Overall composite score. The young adult form includes all of these with the exception of the School Problems composite (Reynolds & Kamphaus, 2004).

For the purposes of this study, the Depression scale, which partially comprises the Internalizing Problems composite of the PRS was analyzed. The Depression scale provides information with regard to feelings of sadness or unhappiness that may adversely affect a child's ability to engage in everyday activities or produce thoughts of suicide (Reynolds & Kamphaus, 2004).

Standardization. The BASC-2 was normed on a sample of 12,850 children between the ages of 2 and 18 years (Reynolds & Kamphaus, 2004). The sample was derived from 256 cities across 40 states and included an equal number of males and females. A stratified, random sampling design was utilized to ensure the inclusion of a representative proportion of children from each demographic group. Data from the March 2001 Current Population Survey was utilized for the basis of stratification along the

following variables: socioeconomic status as measured by parental education level, race/ethnicity, and geographic region. Although data for the normative sample were collected from children in general education settings, the samples included students classified or diagnosed with emotional, behavioral, or physical problems, as reported by their parents (Reynolds & Kamphaus, 2004).

Reliability. The reliability of the BASC-2 is based upon estimates of internal consistency, test-retest reliability, and interrater reliability (Reynolds & Kamphaus, 2004). Internal consistency estimates for the BASC-2 teacher (TRS), parent (PRS), and student (SRP) rating scales yielded median alpha coefficients between .8 and .9. Internal consistency estimates for the PRS Depression and Anxiety were .86 and .82, respectively. Test-retest reliabilities across scales and raters ranged between .65 and .99, indicating moderate to high correlations. The PRS test-retest reliabilities for the Depression and Anxiety scales fell well within this range, with median alpha estimates of .82 and .78, respectively. Although interrater reliability scores for the PRS Depression and Anxiety scales were slightly lower (r = .77 and .69), they reflected a moderate level of reliability (Reynolds & Kamphaus, 2004).

Validity. Evidence for the validity of the BASC-2 was primarily derived from factor analytic studies demonstrating appropriate convergent and divergent validities, as well as correlational research between the BASC-2 and other well-known behavior rating scales (Reynolds & Kamphaus, 2004). With regard to factor analytic studies, correlations within clinical and adaptive scales were found to be positive, whereas correlations

between clinical and adaptive scales were negative. Specifically, the Depression and Anxiety scales were found to be positively correlated with one another, as demonstrated by alpha estimates in the .50 range. With regard to correlational studies, the BASC-2 PRS has been compared against the Achenbach System of Empirically Based Assessment-Child Behavior Checklist (ASEBA-CBCL; Achenbach & Rescorla, 2000; 2001) and the Conners' Parent Rating Scale-Revised (CPRS-R; Conners, 1997). Moderate to high correlations (r = .65 to .75) between the Internalizing scales of the BASC-2 and the ASEBA-CBCL were noted. The correlation between the Anxiety scales of the BASC-2 and the CPRS-R, however, were on the lower end (r = .35 to .41). A review of the scales and items, however, revealed that this discrepancy was expected given that the BASC-2 focuses on general nervousness, fear, and worry, whereas the CPRS-R focuses on emotionality, withdrawal, and timidity (Reynolds & Kamphaus, 2004).

Pediatric Quality of Life Multidimensional Fatigue Scale (PedsQL MFS)

Overview. The Pediatric Quality of Life (PedsQL) Measurement Model (Varni, Seid, & Rode, 1999) is an integrated and modular method to measuring health-related quality of life (HRQOL) in children and adolescents with chronic health conditions (Varni, Burwinkle, Katz, Meeske, & Dickinson, 2002). The PedsQL Measurement Model is grounded in a non-categorical approach that describes children with chronic conditions on the basis of disease-related consequences rather than specific diagnostic labels. While several of the model's measures were derived from data collected as part of a

measurement and prediction study of HRQOL in children with cancer (Varni et al., 1999), they were designed for use across various pediatric populations. The PedsQL Multidimensional Fatigue Scale (MFS) is a recently developed component of the PedsQL Measurement Model designed to capture patient and parent perceptions of fatigue in pediatric patients (Varni et al., 2002).

Content. The PedsQL MFS is an 18-item measure that encompasses three subscales: General Fatigue (6 items), Sleep/Rest Fatigue (6 items), and Cognitive Fatigue (6 items; Varni et al., 2002). The General Fatigue scale incorporates items that tap into a broad sense of fatigue (e.g., I feel tired; I feel too tired to do things that I like to do). The Sleep/Rest Fatigue scale includes items that assess the impact of a medical condition on sleep-related factors (e.g., I feel tired when I wake up in the morning; It is hard for me to sleep through the night). The Cognitive Fatigue scale contains items that measure aspects of mental exhaustion (e.g., It is hard for me to keep my attention on things; It is hard for me to remember what people tell me). Together, these scales produce a Total Fatigue score.

The PedsQL MFS is comprised of parallel child self-report and parent proxy report formats. Child self-report forms are available for children between the ages of 5 and 7 years (young child), 8 and 12 years (child), and 13 to 18 years (adolescent). Parent proxy report forms are available for these age groups in addition to children between the ages of 2 and 4 years (toddler). The items on the forms are identical, differing only in the use of developmentally appropriate language and tense (i.e., first versus third person).

Using a 5-point Likert response scale, the patient or parent is instructed to rate the extent to which each item has been a problem in the past month. Items are reverse scored and linearly transformed into a 0 to 100 point scale, with higher scores indicating fewer problems and thus better HRQOL (Varni et al., 2002).

For the current study, the General Fatigue Self Report score was analyzed. This scale was selected for analysis because it is comprised of items that are identified by pediatric MS patients and their physicians as ones that are commonly experienced and highly relevant to this population (Holland et al., 2012).

Standardization. Normative data for the PedsQL MFS were derived from a sample of healthy children between the ages of 5 and 18 years (n = 52) and parents of children between the ages of 2 and 18 years (n = 102; Varni et al., 2002). Only children between the ages of 5 and 18 years completed the child self-report; whereas, parent proxy reports were completed for all age levels. The sample included reports for 109 males and 45 female, almost half of which were of Hispanic descent (46.1%). The following ethnic groups were also represented: Caucasian (27.5%), African American (5.9%), Asian/Pacific Islander (1.0%), American Indian/Alaskan Native (13.7%), and other (3.9%). Socioeconomic status data were not available (Varni et al., 2002).

Reliability. Evidence for the reliability of the PedsQL MFS was ascertained by calculating Cronbach alpha coefficients to determine internal consistency (Varni et al., 2002). Scores across all child and parent reports exceeded the minimum reliability standard of .70. Additionally, the Total Score across all ages met the .90 criterion, which

is recommend for analyzing individual patient scale scores (Varni et al., 2002). A more recent meta-analytic review noted that the PedsQL was the second most cited instrument for pediatric fatigue with acceptable psychometric properties, including internal consistency, test-retest reliability, and interrater reliability (Tomlinson et al., 2013). Studies involving specific pediatric patient samples, including rheumatology (Varni, Burwinkle, & Szer, 2004), type 1 diabetes (Varni, Limbers, Bryant, & Wilson, 2009), fibromyalgia (Varni, Burwinkle, Limbers, & Szer, 2007), and brain tumors (Palmer, Meeske, Katz, Burwinkle, & Varni, 2007) have all demonstrated adequate internal consistency, with alpha coefficients as high as .94 for the child self-report and 0.94 for the parent proxy report (Varni et al., 2007).

Validity. To ensure the content validity of the PedsQL MFS, multidimensional constructs were derived from an extensive literature review of fatigue in adult and pediatric cancer patients (Varni et al., 2002). Additionally, test developers conducted patient and parent focus groups, pretesting sessions, and field testing trials across two pediatric cancer centers. Construct validity was determined utilizing the known-groups method, in which scale scores are compared across groups known to differ in the construct of interest (i.e., fatigue). Specifically, a comparison between pediatric cancer patients and healthy controls revealed statistically significant differences on self-report and parent-proxy report scores (Varni et al., 2002). The clinical utility and sensitivity of this measure has also been demonstrated in studies involving pediatric rheumotology, fibromyalgia, and diabetes (Varni et al., 2004; 2007; 2009). Concurrent validity is based

on analyses between The PedsQL MFS and two measures: the PedsQL 4.0 Generic Core Scales (Varni et al., 2002) and the Pediatric Functional Assessment of Chronic Illness Therapy-Fatigue (pedFACIT-F; Lai et al., 2007). In the former study, effect sizes were in the moderate to high ranges (.31 to .79; Varni et al., 2002). In the latter study, correlations were reported as .86, .71, and .57 for the general fatigue, sleep fatigue, and cognitive fatigue subscales, respectively (Lai et al., 2007).

Ambulation Index (AI)

Overview. The Ambulation Index (AI; Hauser et al., 1983) is a semiquantitative scale designed to assess changes in gait and mobility by evaluating the time and degree of assistance required to walk 25 feet (Sharrack & Hughes, 1996). The AI was developed with the understanding that a progressive gait dysfunction was an important clinical feature in most patients with MS. Moreover, the authors believed that there was a need to develop a scale that could quantify ambulation-related disability more accurately than existing scoring systems. Consequently, in 1983, the AI was introduced and utilized in a study of 58 adult MS patients receiving intensive immunosuppressive treatments (Hauser et al., 1983).

Content. The AI quantifies ambulation-related disability using an ordinal scale. Scores range from 0 (asymptomatic and fully active) to 9 (wheelchair-bound and unable to transfer independently; Hauser et al., 1983). In order to ascertain a score, the patient is asked to walk a 25-foot course as quickly and safely as possible. The examiner records

the time and type of assistance needed to accomplish the task (Sharrack & Hughes, 1996).

Standardization. Despite its common use in clinical practice, there are currently no standardization data on the AI.

Reliability. Interrater reliability for the AI has been assessed by various sources. Francis, Bain, Swan, and Hughes (1991) reported moderate to high degrees (r = .5 to .7) of interrater agreement and an interrater error rate of 3.9%. The AI was also found to have a greater degree of reproducibility when compared to the EDSS, which had lower agreement scores and higher error rates (Francis et al., 1991). Favorable interrater findings were also reported by Sharrack and colleagues (1999). In this 64-MS patient study, interrater agreement was noted to fall at 77% when agreement was defined as *no difference*, and 100% when defined as a *difference of no more than 1 point*. An overall kappa coefficient of .73 was reported (Sharrack et al., 1999).

Validity. The concurrent validity of the AI is based upon a study (Sharrack et al., 1999) comparing the AI against the following measures: the Barthel Index (Mahoney & Barthel, 1965), the London Handicap Scale (Harwood, Gompertz, & Ebrahim, 1994), the Functional Independence Measure (Keith, Granger, Hamilton, & Sherwin, 1987) and the EDSS (Kurtzke, 1983). Moderate correlations ranging between .68 and .73 were reported (Sharrack et al., 1999).

Data Analysis

Preliminary Analyses

Descriptive statistics. Descriptive statistics were calculated on all independent and dependent variables, including categorical demographic variables (i.e., ethnicity, sex, handedness), continuous demographic variables (i.e., age at baseline, age at follow-up, re-test interval, ambulation), independent clinical variables (i.e., age at onset, disease duration, relapse rate, depression, and fatigue) and all dependent variables (i.e., subtest change scores). Frequencies and percentages for all categorical variables and means, standard deviations, and ranges for all continuous variables were calculated using the Statistical Package for the Social Sciences (SPSS) 21. Additionally, the normality and linearity of subtest scores were investigated using histogram and scatter plot methods, respectively.

Bivariate correlations. The SPSS was also used to conduct bivariate correlations between all variables included in the study. Specifically, independent samples *t* and Mann-Whitney tests were conducted to determine the relationships between categorical and continuous variables, while Spearman's rho correlation coefficients were used to determine associations among continuous variables. Correlation matrices presenting correlation coefficients, significance levels, means, and standard deviations for all variables were generated and examined. These analyses assisted in determining the presence of multicollinearity among variables and in identifying potential control variables (i.e., covariates; Baguley, 2012; Stevens, 2009). Variables characterized by

multicollinearity are highly correlated such that they are redundant when included in statistical analyses (Baguley, 2012). This condition is problematic for a number of statistical operations. In the present study, no variables were found to be so highly correlated as to warrant removal. However, two covariates were identified and included in the regression analysis.

Primary Analyses

Reliable change index (RCI). Choosing the most appropriate method for measuring individual change in performance outcomes over time has been the focus of many psychological investigations (Ronk , Hooke, & Page, 2012; Ronk, Korman, Hooke, & Page, 2013; Walker, Mendella, Stewart, Freedman, & Smith, 2011). Methods for determining individual change should take into consideration whether an individual has made change that is reliable (i.e., a change that is more than what would be expected given the instrument's measurement error) and provide an indication of whether the individual's declines or improvements are clinically meaningful (Ronk et al., 2013). Assessing changes in individual performance on norm-referenced measures, however, can be complicated by several factors, including practice effects, standard error of measurement, and regression toward the mean (Chelune, 2003). While techniques to address these issues have been embedded in several traditional statistical methods used to assess group means, a limited number of reliable techniques have been proposed in the literature to address these issues at the individual level.

The RCI method also referred to as the Jacobson-Truax method, was introduced by Jacobson and Truax (1991) as a means by which to assess change in individual functioning after exposure to psychotherapy, but has since been adopted as an appropriate method for assessing change over time in studies with small sample sizes (Walker et al., 2011). Despite its widespread use, some have argued that the RCI does not account for practice effects and regression toward the mean in the way other statistical methods (i.e., Gulliksen-Lord-Novick method, Edwards-Nunnally method, Hageman-Arrindell method) have purported to do (Chelune, Naugle, Lüders, Sedlak, & Awad, 1993; Hsu, 1995). However, studies (Heaton et al., 2001; Ronk et al., 2012, & Speer, 1992) comparing Jacobson and Truax's RCI against other methods have provided evidence for the RCI's capacity to address regression toward the mean and practice effects with high success (Heaton et al., 2001; Ronk et al., 2012; Speer, 1992).

The RCI is derived from a formula in which the change in a participant's score is divided by the standard error of the difference of the measure being used (Walker et al., 2011). More specifically, the formula used is as follows: RCI = $(T2\text{-}T1)/SE_{diff}$, where *T1* represents the participant's baseline score, *T2* represents the participant's follow-up score, and SE_{diff} represents the standard error of the difference between the two scores. A change is considered significant and reliable if the absolute value of the difference in performance (i.e., change score) exceeds a specified *z*-score cut-off point. For purposes of this study, a *z*-score cut-off value associated with a confidence interval of 90% (i.e., +/- 1.645) was implemented for two reasons. Foremost, the acceptability of this cut-off value

is evidence by its use in prior studies (e.g., Heaton et al., 2001; Walker et al., 2011). Moreover, using a cut-off value associated with a confidence interval of 90%, as opposed to 95%, is recommended in studies with smaller sample sizes so as to increase the likelihood of detecting individuals who fall outside of the 5% interval of the lower and upper ranges of the distribution (Heaton et al., 2001; Walker et al., 2011). A change score was considered to be significant and reliable if it exceeded the *z*-score cutoff value of 1.645 in either a positive or negative direction.

Binomial probability distribution (BPD). Although researchers are often confronted with the task of determining whether an individual presents with cognitive impairment based upon his/her performance across a number of measures, few methods for determining such categorizations have been published (Ingraham & Aiken, 1996), leaving researchers to make subjective decisions based on what they believe to be clinically meaningful. In some cases, researchers have used a specified number of declined measures as a means by which to determine classification into the impaired range, while others have used criterion based on the percent of completed test scores noted to fall below a certain cut-off value (i.e., 1 *SD*, 1.5 *SD*s). Consequently, the literature presents an assortment of criteria, ranging from 2 to 3 measures "impaired" to 33% of the battery "failed" (e.g., Amato et al., 2010; Julian et al., 2013; Walker et al., 2011; Zaaraoui et al., 2011). Interestingly, it is rare for studies to describe the methods used to derive at these criteria.

In determining criteria, researchers should account for the fact that as the number of measures increases, so does the probability that scores on some of the measures will fall within the abnormal or impaired range by virtue of error. To this effect, the BPD method was introduced by Ingraham and Aiken (1996) as a means by which to set an overall criterion for determining impairment when using multiple measures. According to Ingraham and Aiken (1996), the probability of exhibiting at least x_o impaired scores in a battery of *n* tests where the probability of an abnormal score is *t* is derived from:

$$p(x \geq x_o) = 1 - \Sigma c(n, x)t^x(1 - t),^{n-x}$$

Based on this equation, in a battery of 15 tests, a cut-off criterion of two tests at or below 2 standard deviations or three tests at or below 1.5 standard deviations would result in exceeding the criterion less than 5% of the time, thus decreasing the number of false positives (Ingraham & Aiken, 1996). Based on this data, participants in the present study were considered to exhibit significant cognitive impairment if they obtained three or more impaired scores, defined as scores at or below 1.5 standard deviations from the normative mean. This criterion was chosen over the two tests at 2 standard deviations criterion because it is generally consistent with what Amato et al. (2010) used in their longitudinal study of pediatric MS patients. It was believed that maintaining consistency would allow for more viable comparisons.

Repeated measures multivariate analysis of variance (MANOVA). While various methods have been proposed for dealing with measurement of change over time at the group level, repeated measures is the most commonly accepted design for

analyzing longitudinal data (Nakai & Ke, 2009). A repeated measures design has two primary advantages: reduction of variability and recruitment of fewer participants (Baguley, 2012; Stevens, 2009). With regard to the former advantage, because the same participants are measured at baseline and follow-up, they essentially serve as their own controls. Consequently, variability among the participants in each group is eliminated from the error term, which allows for a purer estimate of change. With regard to the latter advantage, because the same participants are being measured repeatedly, fewer participants are required for the study. In other words, a repeated measures design maximizes the value of each participant's contribution, thus eliminating the need to incorporate multiple sets of participants (Baguley, 2012; Stevens, 2009).

Testing for group differences over time using a repeated measures design can be accomplished by employing various types of analyses, including paired samples *t*-test, repeated measures analysis of variance (ANOVA), and repeated measures multivariate analysis of variance (MANOVA; Mertler & Vannatta, 2005). For the purposes of this study, a repeated measures MANOVA was utilized. Repeated measures MANOVA is a statistical procedure designed to test the significance of mean differences among correlated samples (Baguley, 2012; Mertler & Vannatta, 2005). Unlike its univariate counterpart (i.e., repeated measures ANOVA), repeated measures MANOVA is intended for application in studies where several correlated outcome or dependent variables (DV) are being measured as a function of one independent variable (IV; Mertler & Vannatta, 2005; Tabachnick & Fidell, 2013). Theoretically, the DVs are correlated in that they

share an underlying conceptual meaning or framework. Because repeated measures MANOVA treats the multiple DVs in combination, a "composite" DV that maximizes group differences is created from the set of individual DVs. Specifically, the composite DV is developed from a linear equation where all measured DVs have associated weights that are combined and summed to create maximum separation of group means (Tabachnick & Fidell, 2013). In the present study, the composite DV was derived from various individual measures of neurocognitive functioning that are conceivably susceptible to the presence of MS over time.

Repeated measures MANOVA has several advantages over other comparatively simpler univariate methods. First, the use of several fragmented univariate tests results in a greatly inflated overall Type I error rate (Mertler & Vannatta, 2005; Tabachnick & Fidell, 2013). This type of error is likened to a false positive, in which the likelihood of finding spurious and potentially misleading significant results is heightened (Hair, Anderson, Tatham, & Black, 1998; Sherry & Henson, 2005). Using repeated measures MANOVA decreases Type I error by employing a simultaneous analysis procedure that maintains the overall error rate at the pre-selected alpha level (Mertler & Vannatta, 2005; Tabachnick & Fidell, 2013). A second advantage to using repeated measures MANOVA stems from the fact that univariate analyses fail to account for correlations among variables. Repeated measures MANOVA incorporates the intercorrelations among DVs into the analysis, which is essentially the basis for the linear combination of DVs. Lastly, in certain instances, repeated measures MANOVA may reveal differences that do not

emerge in separate repeated measures ANOVAs. For instance, when separate repeated measures ANOVAs are conducted on two DVs, the distribution for each of the groups and DVs might overlap sufficiently, such that a mean difference would not be detected. However, when the two DVs are considered in combination, the two groups may differ substantially, thus resulting in a statistically significant difference between groups. Thus repeated measures MANOVA, which considers all DVs in combination, may occasionally be more powerful than a series of repeated measures ANOVAs (Mertler & Vannatta, 2005; Tabachnick & Fidell, 2013). It is important to note, however, that cases in which MANOVA is more powerful than ANOVA are rare and the use of a more complex statistical analysis, such as a MANOVA, should be carefully considered (Mertler & Vannatta, 2005).

Conducting a repeated measures MANOVA requires three sequential steps including examination of the overall test of significance, analysis of univariate tests on individual DVs, and implementation of post hoc tests (Mertler & Vannatta, 2005). The first step involves testing the tenability of the null hypothesis that all groups are equal on the combination of DVs. This is accomplished by evaluating the significance of the test statistic. While there are several available test statistics for repeated measures MANOVA (e.g., Pillai's Trace, Hotelling's Trace, Roy's Largest Root), the Wilks' Lambda (Λ) criterion was used in the present study, given its extensive use and acceptability in repeated measures MANOVA research. In most cases, if Λ is found to be non-significant, the interpretation of analysis would discontinue and retention of the null hypothesis

would be warranted. Retaining the null hypothesis would indicate that the IV failed to significantly affect the DVs. However, if Λ is significant, the next step would be executed in order to ascertain which of the DVs is affected by the IV. This would be accomplished by conducting a series of univariate ANOVAs on the individual DVs (Mertler & Vannatta, 2005). It is important to note that the present study was interested in examining changes in aggregate performance in addition to performance on various individual assessments. Consequently, univariate ANOVAs were conducted regardless of whether the overall statistic was found to be significant. Moreover, as previously mentioned, ANOVA is often more powerful and thus more sensitive to true significant findings than MANOVA (Mertler & Vannatta, 2005), making this a particularly appropriate approach for this this study.

Running a series of univariate analyses, however, inevitably increases the potential for Type I error. To counteract the potential for inflated error rate, adjustments using Bonferroni correction are typically made to reflect a more stringent alpha level (Mertler & Vannatta, 2005). Bonferroni correction is a common method by which to establish an alpha level for each test so that the overall alpha for the set of DVs does not exceed a specified critical value. Specifically, the critical value for testing each DV is generally the overall alpha level for the analysis divided by the number of comparisons. In cases where studies have a large number of dependent variables, however, correction methods such as Bonferroni are not advisable given the potential for substantially inflating Type II error, making it difficult to detect significant findings and reject the null

hypothesis (Baguley, 2012). For this reasons, a correction method was not implemented in the present study. Finally, a post hoc test should be conducted for any univariate test of a DV that results in significance in order to identify where specific differences lie (Mertler & Vannatta, 2005). In the case of this study, the implementation of post hoc tests were not necessary as only two IV levels (i.e., baseline and follow-up) were analyzed.

Wilcoxon Matched-Pairs Signed-Ranks Test. Conducting analyses in which group means are of interest requires the use of either parametric or non-parametric techniques (Baguley, 2012). Parametric techniques (e.g., *t*-test, ANOVA, MANOVA) assume that data are gathered from a specific probability distribution, and thus model the parameters, or defining properties, of that distribution (Baguley, 2012; Howell, 2007). In other words, parametric tests make inferences or assumptions regarding the distribution of the underlying population from which the sample was culled. Two of the most common parametric assumptions include normality and homogeneity of variance. When these assumptions are met, parametric techniques exhibit great statistical power and precision. However, using parametric techniques can result in erroneous results when the data deviate significantly from the assumptions of the chosen parametric procedure. In such cases, an analogous non-parametric procedure should be utilized (Baguley, 2012; Howell, 2007).

Unlike parametric methods, non-parametric techniques (e.g., Chi-Square Test, Mann-Whitney Test, Wilcoxon Matched-Pairs Signed-Ranks Test) make minimal assumptions with regard to the underlying structure or parameters of the data (Baguley,

2012; Howell, 2007). Because of their robust quality, non-parametric procedures are often used to test the validity of parametric results; such methods are particularly relevant in studies with small sample sizes ($n < 30$) in which there is a lower probability of meeting the assumption of normality. While various non-parametric tests have been noted in the literature, Wilcoxon's Matched-Pairs Signed-Ranks Test is the non-parametric analogue to the *t*-test for related samples and the most commonly used procedure for comparing repeated measurements (Baguley, 2012; Howell, 2007). For these reasons, the Wilcoxon Matched-Pairs Signed-Ranks Test was used in the present study as a means by which to assess the validity of any ANOVA that resulted in significance.

The Wilcoxon Matched-Pairs Signed-Ranks method tests the null hypothesis that the distribution of differences between pairs of scores is equal to zero (Howell, 2007). In order to test this hypothesis, difference scores for each pair must be calculated and ranked, from smallest to largest, without regard to the sign of the difference (i.e., rank order is based on absolute value). All pairs with a difference of zero are ignored, while all pairs with equal absolute differences greater than zero are assigned the same rank value. Once rankings are complete, all difference values are assigned their respective algebraic signs (i.e., negative or positive); all positive and negative ranks are then respectively summed into *W*+ and *W*- to determine the Wilcoxon statistic (*W*). Specifically, *W* is taken as the smaller of the absolute values of the two sums (i.e., *W*+ or *W*-) and is evaluated against the tabled critical values of the Wilcoxon statistic. If *W* is less than the critical

value given for the chosen alpha and number of pairs, the null hypothesis is rejected (Howell, 2007).

Stepwise multiple regression. Various techniques aimed at testing and describing the existence of predictable relationships among a set of variables have been proposed in the literature (e.g., regression analysis, discriminant function analysis, path analysis; Mertler & Vannatta, 2005). Of those techniques, regression analyses have as their primary purpose the development of an equation that can be used for predicting values on some continuous DV for all members of a population, and were thus utilized in the present study. The most basic application of regression analysis is the bivariate or simple linear regression, in which the value of a single DV is predicted from the value of a single IV. In essence, simple regression capitalizes on the correlation between the DV and IV in order to make specific predictions about the DV. The goal of simple regression is to obtain the equation for the best-fitting line (i.e., regression line) through a series of X (IV) and Y (DV) axis points. The extent to which the points are scattered around a line is indicative of the relationship between the IV and the DV, as measured by a correlation coefficient (e.g., Pearson correlation). The higher the correlation coefficient, the higher the degree of predictability between the IV and the DV (Mertler & Vannatta, 2005).

Multiple regression is an extension of the simple regression technique that involves more than one IV or predictor variable (Mertler & Vannatta, 2005). Specifically, this technique is used to predict the value of a single DV from a weighted, linear combination of IVs. One of the most critical aspects of creating the regression model is

selecting a method by which to incorporate the predictor variables into the regression analysis and subsequent equation (Mertler & Vannatta, 2005). Three primary methods have been identified in the literature and include standard multiple regression, sequential (hierarchical) multiple regression, and stepwise (statistical) multiple regression (Tabachnick & Fidell, 2013). In standard multiple regression, all IVs are entered simultaneously into the analysis (Mertler & Vannatta, 2005). The effect of each IV on the DV is assessed as if it has been entered into the equation after all other IVs had been entered. Each IV is then evaluated in terms of what it adds to the prediction of the DV. In sequential multiple regression, IVs are entered into the equation in a specific order. This method is typically chosen when the researcher has substantial knowledge with regard to the potential influence of each variable. The predictor variables hypothesized to have the most influence would be entered into the analysis first (Mertler & Vannatta, 2005).

The third method, stepwise multiple regression, is typically selected for studies that are exploratory in nature (Mertler & Vannatta, 2005). In such cases, the researcher may have a larger set of predictor variables and may want to determine which of the variables make meaningful contributions to the overall predictions. Of note, the sequential and stepwise approaches to regression have one distinct advantage over the standard method; that is, each variable is added at a time and is continually checked across the course of the process for significant improvement to prediction (Mertler & Vannatta, 2005). Considering the purpose of the study and the statistical advantages of the above mentioned methods, stepwise multiple regression was utilized.

In addition to overarching approaches, there are three variations within stepwise multiple regression: forward selection, stepwise selection, and backward deletion (Mertler & Vannatta, 2005). In forward selection, the bivariate correlation among all IVs and the DV are calculated. The IV that has the highest correlation with the DV is then entered into the analysis first and is assessed in terms of its contribution to the DV. The subsequent variable that is entered into the analysis is the IV that contributes most to the DV, after accounting for the effects of the first variable. This process is continued until predictor variables cease to significantly contribute to the prediction of the DV. However, once a variable has been entered, it remains in the analysis. Stepwise selection is an improved variation of forward selection. Using this method, tests are performed at each step to determine the significance of each IV already in the equation. For instance, if a variable entered into the analysis is measuring a similar construct as another variable, the reassessment may determine that the first variable entered may no longer contribute anything to the overall analysis, as its contribution is subsumed by the subsequently inputted variable. In this procedure, the variable that is no longer contributing would be eliminated from the analysis (Mertler & Vannatta, 2005).

In the last variation, backward deletion, an equation including all of the predictor variables is initially computed (Mertler & Vannatta, 2005). Then, a significance test (i.e., partial *F*-test) is conducted on each predictor in order to determine its level of contribution to the overall prediction. The smallest partial *F* is compared to a preselected "*F* to remove" value to determine level of significance. If the value of partial *F* is less

than the "*F* to remove" value, the predictor associated with that partial *F* value is deleted from the analysis and a new equation with the remaining variables is computed. This process is continued until only significant predictors remain in the equation (Mertler & Vannatta, 2005). For purposes of this study, the backward deletion method was utilized to retain variables. This method was chosen because it is preferred in cases wherein a limited number of predictor variables are being analyzed (Kutner, Nachtsheim, Neter, & Li, 2005). However, a reassessment procedure akin to stepwise selection was implemented at each step to ensure the retention of only those variables with the most significant contributions to the model.

Summary

The current study was designed to determine whether pediatric MS patients experience declines in cognitive functioning over a mean period of 14.65 months. Additionally, this study was constructed to elucidate the relationship between cognition and various clinical factors, including age at onset, relapse rate, disease duration, fatigue, and depression. As indicated by the literature review, there is currently a dearth of research related to pediatric MS. Primarily, investigations exploring the evolution of cognitive deficits in MS patients remains rather limited. Additionally, the impact of various clinical factors on cognition remains elusive. Given the potential effects of cognitive functioning on treatment outcomes, educational performance, and overall quality of life, there is a clear need to address these voids. The current study utilized the

RCI and BPD methods, a repeated measures MANOVA, and stepwise multiple regression analyses to accomplish the aforementioned goals.

CHAPTER IV

RESULTS

The primary goal of this study was to determine whether changes in neurocognitive functioning, as measured by performance on specific standardized tests, were evidenced in a sample of pediatric MS patients over a mean period of 14.65 months. The second goal of the study was to examine whether, and to what extent, various clinical factors, including age at onset, disease duration, relapse rate, depression, and fatigue could predict change in neurocognitive functioning over time. The presence of significant changes in neurocognitive performance at the individual level was assessed through use of the Reliable Change Index (RCI) and Binomial Probability Distribution (BPD) methods. Group-based changes in functioning were assessed by examining the overall test of significance produced through a multivariate analysis of variance (MANOVA). The dependent variables most affected by time were located using a series of univariate analyses (i.e., ANOVA) and subsequently confirmed using a nonparametric technique (i.e., Wilcoxon Matched-Pairs Signed-Ranks Test). A series of stepwise multiple regressions were then conducted to determine the predictive relationship between the aforementioned clinical variables of interest and neurocognitive performance across the selected group of measures.

Preliminary Analyses

This section presents descriptive statistics associated with all independent and dependent variables, including demographic and clinical variables of interest. Additionally, bivariate relationships among these variables are reported as either positive or negative. A positive relationship suggests that the variables moved in tandem; whereas, a negative relationship suggests that the variables moved in opposite directions. These analyses assisted in determining the presence of multicollinearity among variables and in identifying potential covariates. Given the study's small sample size and the need to maximize statistical power, the Type I error alpha level for the study was set at $\alpha = .10$. Findings with *p*-values less than .100 are reported as statistically significant. The Statistical Package for the Social Sciences (SPSS) 21.0 was used for all preliminary analyses.

Descriptive Statistics

Frequencies and percentages were calculated for three categorical demographic variables. Information regarding sex, handedness, and ethnicity is reported in Table 1. The sample consisted of 20 pediatric patients with a diagnosis of relapsing-remitting multiple sclerosis (RRMS). The majority of participants in the sample were female (60.0%) and right-handed (90.0%). Almost half of the sample identified themselves as White (45.0%), with Black/African American, Hispanic/Latino, and Asian Indian participants representing the remaining 55.0% of the sample. For the purpose of analysis, participants who identified as White were compared to participants who did not identify

as White due to unequal distributions across ethnic groups. Moreover, handedness and ambulation were utilized for descriptive purposes only. These variables were not considered for further analysis given their unbalanced distributions and limited variations.

Table 1

Frequencies and Percentages for Categorical Demographics Variables

	n	%
Sex		
Female	12	60.0
Male	8	40.0
Handedness		
Right	18	90.0
Left	2	10.0
Ethnicity (Original)		
White	9	45.0
Black	6	30.0
Hispanic	4	20.0
Asian Indian	1	5.0
Ethnicity		
White	9	45.0
Other	11	55.0

Means, standard deviations, and ranges were calculated for all continuous demographic (i.e., age at baseline, age at follow-up, test-retest interval, ambulation) and independent (age at onset, disease duration, relapse rate, depression, fatigue) variables.

Information regarding these variables is displayed in Table 2. Participants in the study were diagnosed with RRMS between the ages of 5 and 16 years (M = 13.70, SD = 2.70) and completed their initial battery between 9 and 18 years of age (M = 14.75, SD = 2.10). Time between baseline and follow-up averaged 14.65 months. At baseline, participants had been diagnosed with RRMS for a mean period of 13.8 months, had experienced an average of 2.25 relapses, and had an average ambulation score (i.e., AI) within the asymptomatic/fully active (M = .25, SD = .64) range.

At follow-up, the means for disease duration and relapse rate had increased to 28.95 months and 2.9 exacerbations, respectively. On the contrary, AI scores were noted to decrease, indicting slight improvements in gait and mobility. On average, participants experienced fatigue at a rate of 67.25 at baseline and 63.65 at follow-up, as measured by the PedsQL MFS General Fatigue Self Report. In the pediatric MS population, "mild" ratings on the PedsQL MFS are defined as between 1 and 2 standard deviations from the mean; whereas, "severe" ratings are defined as 2 or more standard deviations from the mean (MacAllister et al., 2009). Based on normative data from Varni et al. (2002), these scores fall within the mildly impaired range. A mean baseline depression T-score of 55.4, as measured by the BASC-2 PRS, was endorsed. This score falls within the average range of functioning. A follow-up score for depression was not calculated as a result of missing data across three participants. For the purpose of analysis, only baseline scores were utilized in order to assess the extent to which the clinical variables predicted changes in cognitive functioning.

Table 2

Means, Standard Deviations, and Ranges for Continuous Demographic and Independent Variables

	N	*M*	*SD*	Min	Max
Age at Onset	20	13.70	2.70	5	16
Age at Baseline	20	14.75	2.10	9	18
Age at Follow-up	20	16.10	2.13	10	19
Time Between Baseline and Follow-up	20	14.65	3.96	8	28
Ambulation Index (Time 1)	20	.25	.64	0	2
Ambulation Index (Time 2)	20	.15	.49	0	2
Relapse Rate (Time 1)	20	2.25	1.48	1	7
Relapse Rate (Time 2)	20	2.90	1.48	1	7
Disease Duration (Time 1)	20	13.80	12.89	0	43
Disease Duration (Time 2)	20	28.95	15.57	15	68
PedsQL MFS General Self-Report (Time 1)	20	67.25	19.91	38	100
PedsQL MFS General Self-Report (Time 2)	20	63.65	20.31	8	88
BASC-2 Depression Parent Report (Time 1)	20	55.40	18.88	38	120

Note. Age at Baseline and Age at Follow-up are reported in years. Time between Baseline and Follow-up and Disease Duration are reported in months. BASC-2 Depression Parent Report results are represented as *T*-scores ($M = 50$, $SD = 10$).

Means, standard deviations, and ranges were also calculated for all measures utilized in the study. Baseline and follow-up scores associated with each measure are

displayed in Table 3; whereas, the degree of change for each measure is displayed in Table 4. The standardized mean for Digit Span, Symbol Search, and Verbal Fluency (FAS) is 10, with a standard deviation of 3. The standardized mean for CVLT Trials 1-5 is 50, with a standard deviation of 10. The standardized mean for the remaining measures is 100 with a standard deviation of 15. Scores are considered average if they fall within 1 standard deviation from the mean in either a positive or negative direction.

Performance across several measures was found to be higher at baseline than follow-up, suggesting some degree of decline in group performance over time. Performance on Digit Span, for instance, was noted to decline by 1.10 scaled points, the equivalent of approximately five standard points or one-third of a standard deviation. Relatively comparable declines in performance were evidenced across the SDMT and both portions of the Beery (i.e., VMI and VP). In contrast, slight improvements in performance emerged on measures of processing speed (Symbol Search) and verbal fluency (FAS), with more pronounced improvements observed on both portions of the GPT (i.e., Dominant and Non-Dominant). Mixed outcomes were noted across the TMT and the CVLT. Specifically, while mean performance on TMT A was noted to improve by more than one-half of a standard deviation (i.e., 8.15 standard points), performance on TMT B was noted to drop by 4.30 standard points. For the CVLT, half of the scores (i.e., Trials 1-5, Long-Delay Free Recall, Learning Slope) were noted to improve over time, while the remaining half (i.e., Trial 1, Trial 5, CVLT B) were noted to decline.

Table 3

Means, Standard Deviations, and Ranges for Repeated Variables of Interest

	N	*M*	*SD*	Min	Max
Digit Span (Time 1)	20	10.15	3.33	3	15
Digit Span (Time 2)	20	9.05	3.10	4	14
Symbol Search (Time 1)	20	7.95	3.71	1	15
Symbol Search (Time 2)	20	8.75	2.55	3	13
SDMT (Time 1)	20	99.50	20.66	57	134
SDMT (Time 2)	20	95.65	19.90	47	135
CVLT Trials 1–5 (Time 1)	20	45.05	12.43	22	62
CVLT Trials 1–5 (Time 2)	20	46.70	10.44	25	65
CVLT Trial 1 (Time 1)	20	95.75	15.73	70	123
CVLT Trial 1 (Time 2)	20	93.20	15.14	70	130
CVLT Trial 5 (Time 1)	20	96.45	16.05	63	123
CVLT Trial 5 (Time 2)	20	93.05	13.01	70	115
CVLT B (Time 1)	20	90.95	13.92	63	123
CVLT B (Time 2)	20	87.45	14.79	70	130
CVLT Long-Delay Free Recall (Time 1)	20	96.45	13.72	70	115
CVLT Long-Delay Free Recall (Time 2)	20	97.15	15.19	70	123

Table 3, *continued*

Means, Standard Deviations, and Ranges for Repeated Variables of Interest

	N	*M*	*SD*	Min	Max
CVLT Learning Slope (Time 1)	20	96.55	13.19	78	123
CVLT Learning Slope (Time 2)	20	96.95	11.52	78	115
FAS (Time 1)	20	9.85	3.42	4	16
FAS (Time 2)	20	10.35	3.30	5	17
TMT A (Time 1)	20	83.55	24.09	45	124
TMT A (Time 2)	20	91.70	23.37	45	132
TMT B (Time 1)	20	84.30	23.90	45	117
TMT B (Time 2)	20	80.00	25.02	45	118
Beery VMI (Time 1)	20	82.80	9.32	65	100
Beery VMI (Time 2)	20	78.85	14.79	51	103
Beery VP (Time 1)	20	91.90	13.12	45	114
Beery VP (Time 2)	20	88.60	14.67	48	109
GPT Dominant (Time 1)	20	80.85	22.58	45	113
GPT Dominant (Time 2)	20	86.55	23.72	45	119
GPT Non-Dominant (Time 1)	20	81.55	25.65	45	113
GPT Non-Dominant (Time 2)	20	85.20	25.99	45	122

Table 4

Means, Standard Deviations, and Ranges for Dependent Variables

	N	*M*	*SD*	Min	Max
Change in Digit Span	20	-1.10	2.10	-4	5
Change in Symbol Search	20	.80	2.48	-5	5
Change in SDMT	20	-3.85	17.01	-33	29
Change in CVLT Trials 1–5	20	1.65	8.26	-14	16
Change in CVLT Trial 1	20	-2.55	13.61	-23	23
Change in CVLT Trial 5	20	-3.40	12.91	-30	22
Change in CVLT B	20	-3.50	13.14	-30	22
Change in CVLT Long-Delay Free Recall	20	.70	9.09	-22	15
Change in CVLT Learning Slope	20	.40	16.11	-30	30
Change in FAS	20	.50	2.40	-4	4
Change in TMT A	20	8.15	21.15	-45	50
Change in TMT B	20	-4.30	20.65	-57	31
Change in Beery VMI	20	-3.95	11.10	-28	14
Change in Beery VP	20	-3.30	8.86	-19	11
Change in GPT Dominant	20	5.70	17.35	-38	40
Change in GPT Non-Dominant	20	3.65	13.62	-21	31

At baseline, 68.75% of mean scores, with the exception of those for TMT (A and B), Beery VMI, and the GPT (Dominant and Non-Dominant) fell within the average range. At follow-up, 87.5% of mean scores, with the exception of those for TMT B and Beery VMI, were within age expectations. While mean scores indicate average functioning on several measures at baseline and follow-up, there was significant variability among individual performance. Overall, a decrease in mean score performance was evidenced in 50% of the measures.

Bivariate Relationships

Bivariate relationships among demographic variables. Independent samples *t*- and Mann-Whitney tests were conducted to compare the means of female and male participants on age at baseline and time between baseline and follow-up, as depicted in Table 5. Although independent sample *t*-tests failed to reveal a significant effect, the non-parametric Mann-Whitney *U* test revealed a significantly longer period of time between baseline and follow-up for females ($M = 15.75$, $SD = 4.79$) when compared to males ($M = 13.00$, $SD = 1.20$), $p = .038$. Comparable analyses were used to compare the means of White and non-White participants on age at baseline and time between baseline and follow-up; however, no significant relationships were found.

Bivariate relationships among independent variables. Spearman's rho correlations were conducted to examine the relationship among all independent (i.e., clinical) variables of interest, including age at onset, relapse rate, disease duration, fatigue, and depression. Results of this analysis are diaplyed in Table 6.

Table 5

Means and Standard Deviations for Age at Baseline and Time between Baseline and Follow-up by Sex

	N	*M*	*SD*	*t* *U*	*p* *p* (*U*)
Age at Baseline				.21	.835
				55.50	.549
Female	12	14.67	1.61		
Male	8	14.88	2.80		
Time Between Baseline and Follow-up				1.58	.132
				21.50	.038
Female	12	15.75	4.79		
Male	8	13.00	1.20		

Note. Final two columns display test statistics and *p*-values of independent samples *t*-tests and non-parametric Mann-Whitney tests.

As shown in Table 6, results revealed a significant negative correlation between age at onset and disease duration (ρ = -.446), $p < .05$, suggesting that for participants who were diagnosed with RRMS at an older age the course of disease was of shorter duration in comparison to younger participants. Results also revealed a significant positive correlation between relapse rate and disease duration (ρ = .615), $p < .05$, suggesting that participants who had been diagnosed with RRMS for longer periods of time experienced higher frequencies of relapse. A negative, but marginally significant, correlation was also observed between depression and relapse rate (ρ = -.409), $p < .10$.

Table 6

Spearman's Correlations between Independent Variables

	Age at Onset		Relapse Rate		Disease Duration	PedsQL MFS General Self-Report (Time 1)
Age at Onset						
Relapse Rate	-.274					
Disease Duration	-.446	**	.615	***		
PedsQL MFS General Self-Report (Time 1)	.032		.133		.035	
BASC-2 Depression Parent Report (Time 1)	.048		-.409	*	-.283	-.179

Note. * $p < .10$, ** $p < .05$, *** $p < .01$.

Bivariate relationships between demographic and independent variables.

Relationship between age at baseline and time between baseline and follow-up and each of the independent clinical variables were also examined using Spearman's rho correlations. As shown in Table 7, results revealed a significant positive correlation between age at onset and age at baseline ($\rho = .852$), $p < .01$, suggesting that participants who were older at time of diagnosis were also older at baseline. Although bivariate relationships between sex and ethnicity, and each of the independent variables were also examined, no significant relationships were found.

Table 7

Spearman's Rho Correlations between Continuous Independent and Demographic Variables

	Age at Baseline	Time Between Baseline and Follow-up
Age at Onset	.852 ***	-.088
Relapse Rate (Time 1)	-.053	-.299
Disease Duration (Time 1)	-.009	.217
PedsQL MFS General Self-Report (Time 1)	.033	.076
BASC-2 Depression Parent Report (Time 1)	-.084	.054

Note. * $p < .10$, ** $p < .05$, *** $p < .01$.

Bivariate relationships among dependent variables. Spearman's rho was also utilized to examine the relationship among all dependent variables (i.e., change scores) utilized in this study. As shown in Table 8, results revealed a significant positive correlation between the following: change in CVLT Trials 1-5 and change in CVLT Trial 1 ($\rho = .574$), change in CVLT Trial 5 and change in CVLT Long-Delay Free Recall ($\rho = .623$), change in CVLT Learning Slope and change in GPT Dominant ($\rho = .517$), change in FAS and change in Beery VP ($\rho = .572$), and change in Beery VP and change in Beery VMI ($\rho = .461$), $p < .05$. A significant negative correlation, on the other hand, was evidenced between change in Symbol Search and change in Beery VMI ($\rho = -.466$),

between change in CVLT Trial 1 and change in CVLT Learning Slope (ρ = -.649), and between change in CVLT Trial 1 and change in GPT Dominant (ρ = -.456), $p < .05$.

Table 8

Spearman's Rho Correlations between Dependent Variables

	1	2	3	4	5	6	7	8
2	.174							
3	-.133	.343						
4	.047	-.241	-.411					
5	.111	.002	.116	.574 **				
6	-.123	.094	-.048	.421	.152			
7	-.249	.060	.283	.174	.346	.367		
8	.319	.154	.061	.130	.044	.623 **	.208	
9	-.128	.295	-.009	-.369	-.649 **	.437	.033	.219
10	-.393	-.061	.412	-.040	.153	-.183	.256	.085
11	.145	.137	-.028	-.014	-.270	.081	-.169	.040
12	.055	-.168	.077	.038	-.239	.074	-.080	.170
13	.153	-.466 *	.031	-.219	-.150	-.248	-.272	.017
14	-.073	-.370	-.176	.111	.074	-.171	.228	.075
15	-.151	-.035	-.082	-.293	-.456 *	-.188	-.005	-.129
16	-.270	-.153	-.104	-.204	-.022	-.209	.076	-.374

Table 8, *continued*

Spearman's Rho Correlations between Dependent Variables

	9	10	11	12	13	14	15
2							
3							
4							
5							
6							
7							
8							
9							
10	-.371						
11	.269	-.205					
12	.170	-.129	.419				
13	-.118	.111	-.291	-.211			
14	-.417	.572 **	-.285	-.337	.461 *		
15	.517 *	.016	.043	.194	.025	-.152	
16	.129	-.099	-.422	-.122	.212	-.020	.387

Note. * $p < .10$, ** $p < .05$, *** $p < .01$. 1 = Change in Digit Span; 2 = Change in Symbol Search; 3 = Change in SDMT; 4 = Change in CVLT Trials 1–5; 5 = Change in CVLT Trial 1; 6 = Change in CVLT Trial 5; 7 = Change in CVLT B; 8 = Change in CVLT Long-Delay Free Recall; 9 = Change in CVLT Learning Slope; 10 = Change in FAS; 11 = Change in TMT A; 12 = Change in TMT B; 13 = Change in Beery VMI; 14 = Change in Beery VP; 15 = Change in GPT Dominant; 16 = Change in GPT Non-Dominant.

Bivariate relationships between dependent variables and predictors. Lastly, correlations were conducted to examine the relationship between all dependent (i.e.,

change scores) and predictor/control (i.e., clinical and demographic) variables. As shown in Table 9, results revealed a significant positive correlation between change in TMT A and age at onset (ρ = .613) and change in TMT A and age at baseline (ρ = .536), $p < .05$, suggesting that participants who were older at the time of diagnosis and baseline exhibited a greater degree of change in TMT A; more specifically, these participants obtained higher scores on the TMT A at follow-up when compared to baseline. Results also revealed a significant negative correlation between change in GPT Dominant and depression (ρ = -.603), $p < .05$, suggesting that participants with lower depression scores at baseline exhibited a greater degree of change in GPT Dominant; more specifically, these participants obtained higher scores on the GPT Dominant at follow-up when compared to baseline.

As shown in Table 10, independent sample *t*-tests were conducted to compare the means of female and male participants on all dependent variables. Independent samples *t*-tests revealed that females had a significantly higher drop in Digit Span scores (M = -1.75, SD = 1.71) than males (M = -.13, SD = 2.36), although this was not confirmed using the Mann-Whitney test. Moreover, females, on average, had a significantly higher increase in performance on GPT Non-Dominant (M = 8.58, SD = 10.40) in comparison to males who, on average, displayed a decrease in performance on this task (M = -3.75, SD = 15.14), t (18) = 2.17, p = .044. Although comparable analyses were used to assess the bivariate relationships between ethnicity and all dependent variables, no significant relationships were found.

Table 9

Spearman's Rho Correlations between Dependent Variables and Predictors

Change in	1	2	3	4	5	6	7
Digit Span	-.111	-.231	-.114	-.377	-.303	.064	.107
Symbol Search	-.143	.058	-.168	.037	.182	.036	.072
SDMT	-.267	-.393 *	-.112	.114	-.096	-.194	.295
CVLT Trials 1–5	.171	.294	.047	-.169	.114	-.376	.176
CVLT Trial 1	-.268	-.120	-.113	.302	.253	-.356	.285
CVLT Trial 5	.021	.292	.141	-.031	.385 *	-.084	.323
CVLT B	-.383 *	-.198	-.189	.271	.326	-.348	-.128
CVLT Long-Delay Free Recall	-.196	-.005	.281	-.075	.159	.176	.072

Table 9, *continued*

Spearman's Rho Correlations between Dependent Variables and Predictors

Change in	1	2	3	4	5	6	7
CVLT Learning Slope	.000	.125	.001	-.115	.196	.143	-.100
FAS	-.108	-.157	-.145	.393 *	-.040	-.064	-.224
TMT A	.613 **	.536 *	-.034	-.366	-.305	-.094	-.097
TMT B	.293	.107	.171	-.386 *	-.412 *	-.191	.170
Beery VMI	-.037	-.273	.113	-.188	-.358	.275	.074
Beery VP	.108	.052	-.138	.167	-.155	.055	-.288
GPT Dominant	-.109	-.099	-.291	.242	.091	.151	-.603 **
GPT Non-Dominant	-.287	-.262	-.024	.366	.231	-.021	-.247

Note. * $p < .10$, ** $p < .05$, *** $p < .01$. 1 = Age at Onset; 2 = Age at Baseline; 3= Time between Baseline and Follow-up; 4 = Relapse Rate (Time 1); 5 = Disease Duration (Time 1); 6 = PedsQL MFS General Self-Report (Time 1); 7 = BASC-2 Depression Parent Report (Time 1).

Table 10

Means and Standard Deviations for Dependent Variables by Sex

	n	*M*	*SD*	*t* *U*	*p* *p* (*U*)
Change in Digit Span				1.79	.090
				68.50	.105
Female	12	-1.75	1.71		
Male	8	-.13	2.36		
Change in Symbol Search				.29	.777
				55.50	.554
Female	12	.67	2.61		
Male	8	1.00	2.45		
Change in SDMT				.71	.487
				57.00	.487
Female	12	-6.08	16.16		
Male	8	-.50	18.81		
Change in CVLT Trials 1–5				.33	.742
				46.50	.908
Female	12	2.17	8.84		
Male	8	.88	7.83		
Change in CVLT Trial 1				.95	.355
				59.50	.369
Female	12	-4.92	12.32		
Male	8	1.00	15.51		
Change in CVLT Trial 5				.10	.924
				50.50	.845
Female	12	-3.17	11.64		
Male	8	-3.75	15.46		
Change in CVLT B				.17	.868
				50.50	.845
Female	12	-3.92	15.08		
Male	8	-2.88	10.49		

Table 10, *continued*

Means and Standard Deviations for Dependent Variables by Sex

	n	*M*	*SD*	*t* *U*	*p* *p* (*U*)
Change in CVLT Long-Delay Free Recall				.51 55.50	.615 .555
Female	12	-.17	10.05		
Male	8	2.00	7.87		
Change in CVLT Learning Slope				.25 43.50	.802 .727
Female	12	1.17	18.33		
Male	8	-.75	13.18		
Change in FAS				.56 40.00	.582 .532
Female	12	.75	2.56		
Male	8	.13	2.23		
Change in TMT A				1.10 65.00	.285 .189
Female	12	3.92	23.57		
Male	8	14.50	16.27		
Change in TMT B				.60 37.00	.556 .396
Female	12	-2.00	24.19		
Male	8	-7.75	14.67		
Change in Beery VMI				.06 49.50	.956 .908
Female	12	-3.83	11.20		
Male	8	-4.13	11.70		
Change in Beery VP				.07 45.50	.945 .847
Female	12	-3.42	7.94		
Male	8	-3.13	10.68		

Table 10, *continued*

Means and Standard Deviations for Dependent Variables by Sex

	n	*M*	*SD*	*t* *U*	*p* *p* (*U*)
Change in GPT Dominant				.17	.871
				45.00	.817
Female	12	6.33	10.28		
Male	8	4.75	25.49		
Change in GPT Non-Dominant				2.17	.044
				22.50	.049
Female	12	8.58	10.40		
Male	8	-3.75	15.14		

Note. Final two columns display test statistics and *p*-values of independent samples *t*-tests and non-parametric Mann-Whitney tests.

Primary Analyses

This section presents the primary analyses that were used to address the research questions and their corresponding hypotheses. Findings with *p*-values less than .100 are reported as statistically significant. The Statistical Package for the Social Sciences (SPSS) 21.0 was used for all primary analyses.

Reliable Change Index and Binomial Probability Distribution

The RCI method (Jacobson & Truax, 1991) was used to determine the clinical significance of individual differences in performance over time. Specifically, RCI calculations were used to classify performance scores for each of the 20 participants into one of three categories: Declined, Stable, or Improved. The cut-off value for the

classification of the RCI scores were set at +/- 1.645, such that all change scores below -1.645 were classified as Declined, scores between -1.645 and 1.645 were classified as Stable, and scores above 1.645 were classified as Improved. Table 11 displays the number of tests classified into these categories for each of the participants.

Table 11

Frequency of Declined, Stable, and Improved Measures per Participant using the Reliable Change Index (RCI) Method and Cognitive Status of Participant at Baseline and Follow-up using the Binomial Probability Distribution (BPD) Method

Participant	Number of Declined Measures	Number of Stable Measures	Number of Improved Measures	Status
1	8	4	4	CI/CI
2	5	5	6	CI/CI
3	6	1	9	S/CI
4	7	3	6	S/S
5	5	5	6	S/S
6	2	7	7	S/S
7	7	2	7	S/S
8	6	4	6	S/S
9	5	4	7	S/S
10	5	4	7	CI/CI
11	5	5	6	CI/CI
12	9	3	4	S/S
13	9	2	5	S/S
14	3	7	6	S/S
15	6	3	7	S/S
16	5	3	8	CI/S
17	7	6	3	CI/CI
18	2	7	7	CI/CI
19	6	4	6	CI/CI
20	11	3	2	CI/CI

Note. CI = Cognitively Impaired; S = Stable

Thirty percent of participants showed more declines than they did improvements; whereas, 55% of participants showed more improvements than they did declines. All remaining participants (15%) had an equal number of improvements and declines. The frequencies of improvement and decline on each measure are displayed in Table 12. Declines were primarily evidenced on measures of auditory working memory (Digit Span; 65%), visual tracking and sustained visual attention (SDMT; 60%), immediate verbal learning (CVLT Trial 1; 50%), and visual perception (Beery VP; 50%). Conversely, participants improved primarily on measures of motor-based processing speed (Symbol Search; 50%), verbal fluency (FAS; 60%), and manual speed and dexterity (GPT Non Dominant; 50%). GPT Dominant was noted as being the measure on which participants exhibited the most stability, with TMT A as the second most stable measure.

Significant cognitive impairment was identified using the BPD method. Specifically, participants with at least three scores at or below 1.5 standard deviations from the normative mean were classified as exhibiting significant cognitive impairment at both baseline and follow-up. All other participants were categorized as stable. As presented in Table 11, 40% of participants met criteria for cognitive impairment at both time points; whereas, 50% of particpants were noted to remain stable across time. One participant who was noted to meet criteria for cognitive impairment at baseline failed to meet criteria at follow-up, while another participant who had failed to meet criteria for cognitive impairment at baseline was categorized as such at follow-up.

Table 12

Frequency and Percent of Participants Exhibiting Significant Cognitive Change per Measure Using the Reliable Change Index (RCI) Method

	Declined	Stable	Improved	RCI
Digit Span	13 (65)	3 (15)	4 (20)	-2.34
Symbol Search	4 (20)	6 (30)	10 (50)	1.44
SDMT	12 (60)	2 (10)	6 (30)	-1.01
CVLT Trials 1–5	5 (25)	6 (30)	9 (45)	0.89
CVLT Trial 1	10 (50)	5 (25)	5 (25)	-0.84
CVLT Trial 5	9 (45)	5 (25)	6 (30)	-1.18
CVLT B	8 (40)	5 (25)	7 (35)	-1.19
CVLT Long-Delay Free Recall	6 (30)	6 (30)	8 (40)	0.34
CVLT Learning Slope	8 (40)	(15)	9 (45)	0.11
FAS	6 (30)	2 (10)	12 (60)	0.93
TMT A	3 (15)	8 (40)	9 (45)	1.72
TMT B	8 (40)	6 (30)	6 (30)	-0.93
Beery VMI	9 (45)	5 (25)	6 (30)	-1.59
Beery VP	10 (50)	6 (30)	4 (20)	-1.67
GPT Dominant	3 (15)	9 (45)	8 (40)	1.47
GPT Non-Dominant	5 (25)	5 (20)	10 (50)	1.20

Note. The last column displays average RCI values for each measure.

Repeated Measures Multivariate Analysis of Variance

A MANOVA was conducted to test for mean differences in neurocognitive performance over time. Multivariate results indicated a significant overall effect, F (16, 4) = 10.05, p = .019, η^2 = .976, as depicted in Table 13. Examination of the univariate effects revealed a significant difference on Digit Span, F (1, 12.10) = 5.49, p = .030, η^2 = .224, such that particpants scores at baseline were significantly higher (M = 10.15, SD = 3.33) than at follow-up (M = 9.05, SD = 3.10). Results were verified using the Wilcoxon Matched-Pairs Signed-Ranks Test, W = 26.00, p = .016. This test also revealed a significant difference on TMT A scores from baseline to follow-up, W = 129.50, p = .055. On average, participants scored higher on this measure at follow-up (M = 91.70, SD = 23.37) when compared to baseline (M = 83.55, SD = 24.09).

Table 13

Means and Standard Deviations for Baseline and Follow-up Measures from MANOVA

	N	*M*	*SD*	*F* *W*	*p* *p* (*W*)
Digit Span				5.49 26.00	.030 .016
Time 1	20	10.15	3.33		
Time 2	20	9.05	3.10		
Symbol Search				2.08 76.50	.166 .129
Time 1	20	7.95	3.71		
Time 2	20	8.75	2.55		
SDMT				1.02 76.00	.324 .278
Time 1	20	99.50	20.66		
Time 2	20	95.65	19.90		

Table 13, *continued*

Means and Standard Deviations for Baseline and Follow-up Measures from MANOVA

	N	*M*	*SD*	*F* *W*	*p* *p* (*W*)
CVLT Trials 1–5				.80	.383
				120.00	.314
Time 1	20	45.05	12.43		
Time 2	20	46.70	10.44		
CVLT Trial 1				.70	.413
				48.50	.510
Time 1	20	95.75	15.73		
Time 2	20	93.20	15.14		
CVLT Trial 5				1.39	.253
				38.50	.219
Time 1	20	96.45	16.05		
Time 2	20	93.05	13.01		
CVLT B				1.42	.248
				35.50	.162
Time 1	20	90.95	13.92		
Time 2	20	87.45	14.79		
CVLT Long-Delay Free Recall				.12	.734
				56.50	.798
Time 1	20	96.45	13.72		
Time 2	20	97.15	15.19		
CVLT Learning Slope				.01	.913
				78.00	.943
Time 1	20	96.55	13.19		
Time 2	20	96.95	11.52		
FAS				.87	.362
				104.00	.416
Time 1	20	9.85	3.42		
Time 2	20	10.35	3.30		
TMT A				2.97	.101
				129.50	.055
Time 1	20	83.55	24.09		
Time 2	20	91.70	23.37		

Table 13, *continued*

Means and Standard Deviations for Baseline and Follow-up Measures from MANOVA

	N	*M*	*SD*	*F* *W*	*p* *p* (*W*)
TMT B				.87	.363
				66.00	.396
Time 1	20	84.30	23.90		
Time 2	20	80.00	25.02		
Beery VMI				2.54	.128
				61.00	.171
Time 1	20	82.80	9.32		
Time 2	20	78.85	14.79		
Beery VP				2.77	.112
				51.00	.133
Time 1	20	91.90	13.12		
Time 2	20	88.60	14.67		
GPT Dominant				2.16	.158
				135.00	.107
Time 1	20	80.85	22.58		
Time 2	20	86.55	23.72		
GPT Non-Dominant				1.44	.245
				112.00	.248
Time 1	20	81.55	25.65		
Time 2	20	85.20	25.99		

Note. Multivariate F (16, 4) = 10.05, p = .019, partial η^2 = .976. Final two columns display test statistics and p-values of paired samples t-tests and non-parametric Wilcoxon Matched-Pairs Signed-Ranks tests.

Stepwise Multiple Regression

A series of stepwise multiple regressions were conducted to predict change in each of the dependent measures using age at onset, relapse rate, disease duration, fatigue, and depression. Variables identified in the preliminary analyses as having the potential to affect study outcomes were included as control variables (i.e., covariates). More specifically, sex was considered as a potential covariate in the models predicting change

in Digit Span and GPT Non-Dominant and age at baseline was considered as a potential covariate for the models predicting change in SDMT and TMT A. Summaries of all models are displayed in Tables 14-17. Given the limited number of predictor variables in the study, backward deletion was used to arrive at the model with the highest adjusted R-squared. However, a reassessment procedure akin to stepwise selection was implemented at each step to ensure the inclusion of only those variables with the most significant contributions to the model.

Summaries of regression models predicting changes in performance on Digit Span, Symbol Search, SDMT, and CVLT Trials 1-5 are displayed in Table 14. Of these four models, two were found to be statistically significant. The overall model predicting change in Digit Span was significant, F (3, 16) = 4.06, p = .025, with sex, relapse rate, and depression accounting for 43.2% of the total variance. In this model, sex and depression were both significant predictors (ps < .10). Specifically, male participants and participants with higher levels of depression at baseline were more likely to exhibit an increase in performance on Digit Span over time (*Beta* = .384 and .381 respectively, ps < .10). The overall model predicting change in SDMT was also significant, F (3, 16) = 3.81, p = .031, with age at onset, disease duration, and depression accounting for 41.7% of the variance. Age at onset was the only significant predictor in this model (p < .05). Specifically, as age at onset increased, change scores on SDMT decreased (*Beta* = -.691, p < .01); that is, participants who were older at time of diagnosis were more likely to experience decreases in SDMT performance over time. Although the overall model for

change in CVLT Trials 1-5 was not significant, higher levels of fatigue at baseline predicted a decrease in performance on this task over time (*Beta* = -.466, $p < .05$).

Table 14

Multiple Linear Regression Predicting Change in Digit Span, Change in Symbol Search, Change in SDMT, and Change in CVLT Trials 1–5

	Digit Span	Symbol Search	SDMT	CVLT Trials 1–5
Sex (Male)	.384 *	—	—	—
Age at Onset	—	—	-.691 ***	.343
Relapse Rate (Time 1)	-.296	-.340	—	—
Disease Duration (Time 1)		.319	-.300	.255
PedsQL MFS General Self-Report (Time 1)	—	—	—	-.466 **
BASC-2 Depression Parent Report (Time 1)	.381 *	.233	.306	—
F (3, 16)	4.06 **	1.31	3.81 **	2.29
R^2	.432	.197	.417	.300
Adj. R^2	.325	.047	.307	.169

Note. * $p < .10$, ** $p < .05$, *** $p < .01$. F-test degrees of freedom are shown in parentheses.

Summaries of regression models predicting changes in performance on CVLT Trial 1, CVLT Trial 5, CVLT B, and CVLT Long-Delay Free Recall are displayed in Table 15. Of these four models, three were found to be statistically significant. The overall model predicting change in CVLT Trial 1 was significant, F (4, 15) = 4.50, p = .014, with relapse rate and depression accounting for 54.5% of the total variance.

Table 15

Multiple Linear Regression Predicting Change in CVLT Trial 1, Change in CVLT Trial 5, Change in CVLT B, and Change in CVLT Long-Delay Free Recall

	CVLT Trial 1		CVLT Trial 5		CVLT B		CVLT Long Delay Free Recall
Age at Onset	-.233		.266		—		—
Relapse Rate (Time 1)	.475	**	—		—		—
Disease Duration (Time 1)	—		.598	**	.439	*	.142
PedsQL MFS General Self-Report (Time 1)	-.216		—		—		—
BASC-2 Depression Parent Report (Time 1)	.461	**	.300		—		—
F	4.50 (4, 15)	**	2.46 (3, 16)	*	4.30 (1, 18)	*	0.37 (1, 18)
R^2	.545		.316		.193		.020
Adj. R^2	.424		.187		.148		-.034

Note. * $p < .10$, ** $p < .05$, *** $p < .01$. *F*-test degrees of freedom are shown in parentheses.

Results from this model indicate that higher numbers of relapse and higher levels of depression at baseline were predictive of increases in performance on CVLT Trial 1 (*Beta* = .475 and .461, *p*s < .05). The overall models predicting change in CVLT Trial 5 and CVLT B were also significant, F (3, 16) = 2.46, p = .100; F (1, 18) = 4.30, p = .053. In these models, longer disease duration was predictive of increases in CVLT Trial 5 and CVLT B performance, (*Beta* = .598, $p < .05$; *Beta* = .439, $p < .10$). While all three models were found to be statistically significant, only the model for CVLT Trial 1 displayed a moderate predictive power (R^2 = .545). The other two models were unable to

explain a large proportion of the variance in their respective dependent variables (R^2s < .350).

Summaries of regression models predicting changes in performance on CVLT Learning Slope, FAS, TMT A, and TMT B are displayed in Table 16. Of these four models, only the model predicting change in TMT A was found to be significant, F (4, 15) = 2.75, p = .067. In this model, higher levels of fatigue at baseline were predictive of a decrease in performance on TMT A over time (*Beta* = -.372, p < .10).

Table 16

Multiple Linear Regression Predicting Change in CVLT Learning Slope, Change in FAS, Change in TMT A, and Change in TMT B

	CVLT Learning Slope	FAS	TMT A	TMT B
Age at Onset	—	-.292	—	—
Age at Baseline	—	—	.323	—
Relapse Rate (Time 1)	-.348	.313	-.233	—
Disease Duration (Time 1)	.351	-.398	-.302	-.275
PedsQL MFS General Self-Report (Time 1)	—	—	-.372 *	-.250
BASC-2 Depression Parent Report (Time 1)	—	-.274	—	—
F	1.55 (2, 17)	1.52 (4, 15)	2.75 * (4, 15)	1.09 (2, 17)
R^2	.154	.288	.423	.114
Adj. R^2	.055	.098	.269	.009

Note. * p < .10, ** p < .05, *** p < .01. F-test degrees of freedom are shown in parentheses.

Finally, summaries of regression models predicting changes in performance on Beery VMI, Beery VP, GPT Dominant, and GPT Non-Dominant are displayed in Table 17. Of the four models, the ones predicting changes in Beery VP and GPT Dominant were statistically significant, F (2, 17) = 2.84, p = .087; F (1, 18) = 22.24, p = .000. In both models, higher depression scores at baseline were predictive of decreases in task performance over time (*Beta* = -.477, $p < .05$; *Beta* = -.743, $p < .01$).

Table 17

Multiple Linear Regression Predicting Change in Beery VMI, Change in Beery VP, Change in GPT Dominant, and Change in GPT Non-Dominant

	Beery VMI	Beery VP	GPT Dominant	GPT Non-Dominant
Age at Onset	-.315	—	—	—
Age at Baseline	—	—	—	-.183
Relapse Rate (Time 1)	—	—	—	.238
Disease Duration (Time 1)	-.443	-.264	—	.122
BASC-2 Depression Parent Report (Time 1)	—	-.477 **	-.743 ***	—
F	1.55 (2, 17)	2.84 * (2, 17)	22.24 *** (1, 18)	.904 (3, 16)
R^2	.154	.250	.553	.145
Adj. R^2	.055	.162	.528	-.015

Note. * $p < .10$, ** $p < .05$, *** $p < .01$. F-test degrees of freedom are shown in parentheses.

Summary

The manifestation of change in neurocognitive functioning over time in a sample of pediatric MS patients was first assessed using individual outcome procedures, which indicated clinically significant declines across participants, particularly in the areas of auditory working memory (Digit Span), visual tracking and sustained visual attention (SDMT), immediate verbal learning (CVLT Trial 1), and visual perception (Beery VP). Additionally, a multivariate procedure conducted to assess group changes resulted in an overall significant effect. Subsequent univariate and non-parametric analyses revealed significant differences in mean performance on Digit Span and TMT A. Specifically, while Digit Span scores were noted to decline from baseline to follow-up, TMT A scores were noted to increase over time. A series of regressions were then employed to predict group-based changes in neurocognitive functioning using age at onset, relapse rate, disease duration, fatigue, and depression in addition to sex and age at baseline, which were identified as potential covariates in the preliminary analyses. Results revealed a significant model for working memory, as measured by Digit Span, with sex (i.e., male) and higher rates of depression at baseline as predictors of improved performance over time. Improvements in performance were also evidenced on various aspects of memory (i.e., CVLT Trial 1, CVLT Trial 5, and CVLT B) as a function of greater relapse rates, higher levels of depression, and longer periods of disease duration. Results also revealed a significant model for visual scanning/tracking and sustained visual attention, as measured by the SDMT, with older age at onset as a predictor of decreases in

performance over time. Decreases in performance were also evidenced on a task of visual scanning and motor speed (i.e., TMT A) as a function of higher levels of fatigue at baseline and on tasks of visual perception and manual speed and dexterity (i.e., Beery VP, GPT Dominant) as a function of higher depression scores at baseline. Overall, results suggest that pediatric MS patients experience statistically and clinically meaningful changes in neurocognitive performance over time that are likely predicted by various relevant clinical and demographic variables.

CHAPTER V

DISCUSSION

The present study assessed cognitive functioning in a cohort of pediatric MS patients using individual and group-level methodologies. Based upon a review of the literature suggesting the emergence of significant cognitive impairment across time (i.e., 2 years), it was hypothesized that participants in this study would demonstrate declines in verbal fluency, complex attention (i.e., sustained, shifting), and verbal learning and memory across a relatively short period of time (i.e., 14.65 months). The study was also conducted in order to elucidate the predictive relationship between cognitive functioning and various clinical factors, including age at onset, relapse rate, disease duration, fatigue, and depression. Specifically, the study assessed whether, and to what extent, changes in neurocognitive performance correlated with the aforementioned factors using stepwise multiple regression analyses. It was hypothesized that depression and fatigue would be negatively correlated with cognitive performance over time; whereas, a positive correlation between age at onset and cognitive functioning would be evidenced.

Depression was selected as a viable predictor based on prior research suggesting an inverse relationship between depressive symptomology and planning ability, working memory capacity, and attentional processes (Julian & Arnett, 2009; Julian et al., 2011). The predictive power of fatigue was derived from its complex pathophysiology and impact on the central nervous, immune, and neuroendocrine systems (MacAllister &

Krupp, 2005). Lastly, younger age at onset was selected based on theories of development. Specifically, younger age has been associated with the presence of immature pathways and developing networks that are highly susceptible to the pathogenic processes of MS (Fields, 2008; Julian et al., 2011).

Explanation of Findings

Sample Characteristics

With regard to demographic features and disease characteristics, the composition of the study's sample was found to be generally consistent with other pediatric MS studies. All participants, for instance, presented with a relapsing-remitting course of disease. This was expected given that RRMS accounts for 85.7 to 100% of all pediatric MS cases (Waubant & Chabas, 2009). The 3:2 female to male ratio of the sample is also consistent with literature suggesting a higher preponderance of affected females, particularly in cohorts with a large representation of children between the ages of 6 and 18 years (Chitnis et al., 2009; Simone et al., 2002). The ethnic distribution of the sample was also extremely diverse (i.e., higher proportion of non-White participants) in comparison to what is typically seen in adult studies. This finding might simply reflect the changing demographics of the United States and the proximal catchment area of the hospital, but it might also speak to environmental risk factors specific to minority children. This supposition is quite tentative and merely highlights the need for additional epidemiological investigations. In any case, findings related to the ethnic distribution of the sample were consistent with studies suggesting a higher proportion of African

American and Hispanic individuals in pediatric-onset group studies (Banwell et al., 2009; Krupp, McLinskey, & Troell, 2008).

As expected, the distribution for age at onset in the sample was not characterized by the bell shaped pattern typically observed in the overall MS population (Ruggieri et al., 2011) given the exclusion of adult participants from the study. Nonetheless, the sample closely approximated the expected distribution for age at onset in the pediatric population, with the majority of participants (90%) diagnosed between the ages of 12 and 16 years. With regard to exacerbations, the sample approximated the estimated annual relapse rate of .38 to .87 (Renoux & Waubant, 2011). Specifically, the average rate of relapse in the 14.65 months between baseline and follow-up was .65. The overall extent of self-reported fatigue in this sample was noted to fall in the mildly elevated range. An elevated level of fatigue is not unusual given the impact of MS on the immune, central nervous, and neuroendocrine systems, all of which contribute to fatigue-related symptomology (MacAllister & Krupp, 2005).

With that said, the reported rates of fatigue in this sample were noted to be higher than those reported in other studies (Amato, Zipoli, et al., 2008; MacAllister et al., 2009), with 80 and 90% of participants reporting either mild or severe symptoms at baseline and follow-up, respectively. These findings may suggest that the rate of fatigue in pediatric patients approximates those of adult patients more closely than had been previously reported. However, research related to the incidence and severity of fatigue among pediatric cohorts is not well-established and these results should be used to formulate

hypotheses rather than definitive conclusions. The mean rate of depression in this sample was noted to fall within the average range though individual scores varied widely. Analysis of individual depression scores at baseline revealed that at least 35% of the sample exhibited elevated levels of depression. This finding is consistent with evidence suggesting the presence of significant psychological distress (i.e., depression, anxiety) among pediatric cohorts (Amato, Zipoli, et al., 2008; MacAllister, Boyd, et al., 2007).

Descriptive analyses of the sample's collective performance across measures utilized in the current study revealed that participants experienced increases and declines at equal rates. That is, participants' mean scores were noted to increase on half of the measures and decrease on the remaining half. Scores were noted to increase on tasks associated with motor-based processing speed (Symbol Search), verbal fluency (FAS), visual scanning and motor speed (TMT A), and manual speed and dexterity (GPT Dominant and Non-Dominant). Increments in performance were also observed on a global measure of immediate verbal learning (CVLT Trials 1-5), a measure of delayed verbal learning (CVLT Long-Delay Free Recall), and a measure of verbal learning threshold (i.e., CVLT Learning Slope). In contrast, scores were noted to decrease on tasks of working memory and mental capacity (Digit Span), visual tracking and sustained visual attention (SDMT), shifting attention (TMT B), motor coordination and spatial organization (Beery VMI), and visual perception (Beery VP). Decrements in performance were also observed on three tasks assessing the immediate recall of verbal information (i.e., CVLT Trial 1, CVLT Trial 5, and CVLT Trial B).

Although previous longitudinal studies have focused primarily on the frequency of participants impaired at baseline in comparison to follow-up, information regarding group-based changes was presented in a study by MacAllister, Christodoulou, and colleagues (2007). In this study, participants (n = 12) were noted to exhibit increases in performance on measures of immediate and delayed verbal learning (WRAML) and declines on measures of verbal fluency (COWAT), visual scanning and motor speed (TMT A), shifting attention (TMT B), and motor coordination and spatial organization (Beery VMI). To some extent, the results of the current study are consistent with these findings. Specifically, both samples exhibited improvements on a global measure of immediate verbal learning. Of note, the sample in the current study experienced decreases on two of the trials (i.e., Trials 1 and 5) that comprise the global measure (i.e., CVLT Trials 1-5). Although the careful examination of learning strategies and recall errors was not feasible in this case, it is plausible that the increase in the global measure was reflective of adequate performance on Trials 2 through 4, as opposed to Trials 1 and 5. Symptoms commonly experienced by pediatric MS patients, including fatigue and compromised attention, could account for adequate performance on some trials as opposed to others.

Both samples also experienced increases on a measure of delayed verbal learning. An examination of the follow-up data revealed that, on average, participants maintained the same level of performance at Trial 5 that they exhibited at Trial 1. However, participants exhibited a slight improvement between the immediate and delayed recall

portions of the task. Taken together, this suggests that children with MS may benefit, to some extent, from a period of consolidation after repeated exposure to verbal information. This is expected given the effects of demyelination on speeded acquisition of information, which supports learning (Fields, 2008; Nagy et al., 2004).

With regard to mean decreases in performance, both samples experienced declines on tasks of motor coordination (Beery VMI) and shifting attention (TMT B). With regard to the former finding, it is important to consider the role of white matter integrity in the functional connectivity of distant brain regions, which supports the integration of neuropsychological processes and functions (Fields, 2008). Because the Beery VMI is considered to be a task of integration (i.e., requires the integration of functions), performance on this task is likely to be affected by white matter disruption, such as that which is experienced by patients with MS. With regard to the latter findings, the TMT B has consistently demonstrated sensitivity to neurological impairment in a number of populations, including pediatric MS (Amato et al., 2010; MacAllister, Christodoulou, et al., 2007). Moreover, the prominent difficulties in complex attention and executive functioning in pediatric patients has been substantiated in cross-sectional and longitudinal studies (Amato, Gorretti, et al., 2008; Portaccio et al. 2009; MacAlliser, Christodoulou et al., 2007). Additional declines in performance were evidenced in the areas of auditory working memory (Digit Span), visual tracking and sustained visual attention (SDMT), and visual perception (VP). While prior longitudinal studies have not descriptively addressed group changes across these domains, these findings substantiate

literature pertaining to the potential adverse impact of demyelination on attention, memory, and executive functioning.

Outcomes related to performance on measures of verbal fluency (FAS), processing speed (Symbol Search), and visual scanning and motor speed (TMT A) are particularly noeworthy. Although studies have demonstrated relatively intact performance on measures of expressive language in adult MS patients, decreases in performance in verbal fluency have been evidenced in pediatric cohorts (Amato et al., 2010; MacAllister, Christodoulou, et al., 2007). Results from this study, however, indicate an increase in mean performance over time, with 80 to 85% of participants functioning within the average range at baseline and follow-up, respectively. Differences in performance between the current study's sample and other longitudinal samples might be partially attributable to the demands of the tasks utilized in the studies. Specifically, Amato et al. (2010) used phonemic (initial letter prompt) as well as a semantic (categories) verbal fluency tasks. An analysis of the data provided in the aforementioned study indicated a higher rate of impairment on the semantic task in comparison to the phonemic task. With that said, it is difficult to compare aggregate changes in performance given that the Amato et al. (2010) study did not report this data.

The findings suggesting positive mean changes over time on Symbol Search and TMT A are also noteworthy. With regard to the former finding, increases in performance were not expected based on what is known about the direct impact of demyelination on speed of information processing. With regard to the latter finding, while pediatric MS

patients have consistently demonstrated difficulties on the TMT A, direct comparisons between change scores can only be made between this study and that of MacAllister, Christodoulou, et al. (2007). In comparing the two studies, there is a clear discrepancy between outcomes, with participants from the current study exhibiting a mean change score of 8.15 and participants from the 2007 study exhibiting a mean change score of -11.08. While it is difficult to provide conclusive explanations for this discrepancy, there are numerous factors, such as resiliency, treatment outcomes, and cognitive reserve that should be considered when interpreting data related to changes in cognitive functioning. The potential impact of these factors are discussed in the limitations and considerations section.

Relationships between Variables of Interest

Univariate analyses focused upon demographic variables indicated a relationship between sex and test-retest interval such that females had a significantly longer period of time between baseline and follow-up when compared to males. Upon examination of the data, it was found that one female had a test-retest period of 28 months, which is regarded as an outlier when considering the modes (12, 13), mean (14.65), and median (14) of the sample. In this case, correlational findings were significantly influenced by the presence of an outlier.

Correlational analyses among all independent variables in the study revealed a significant negative correlation between age at onset and disease duration, such that participants who were diagnosed with RRMS at an older age exhibited a shorter course of

disease in comparison to younger participants. Because disease duration is the residual between a patient's age at testing and age at onset, this finding is unremarkable. Results also revealed a significant positive correlation between relapse rate and disease duration, such that participants who had been diagnosed with RRMS for longer periods of time experienced higher frequencies of relapse. Although the course of RRMS can be largely heterogeneous and dependent upon a patient's response to treatment, it is estimated that children experience between .38 and .87 relapses annually (Renoux & Waubant, 2011). Thus, children with longer exposure to RRMS (i.e., longer disease durations) are at a higher risk for experiencing more relapses when compared to children with briefer disease durations.

Significant relationships also emerged between all of the dependent variables. The fact that all measures exhibited some degree of relationship with one another follows theoretical reasoning. Assuming that neurocognitive processes are physiologically and functionally connected to some extent (Colom et al., 2003; Gazzaniga, Ivry, Mangun, & Steven, 1998), and that the measures in the study were developed to reliably assess various aspects of neurocognitive functioning, one would expect some degree of relationship to emerge. The directionality and interpretability of these relationships, however, varied. Almost all tasks from the CVLT were noted to positively correlate with one another. This finding, however, is excpected given the computational links between some of the tasks (Delis et al., 1994; 2000). For instance, CVLT Trial 1 directly contributes to CVLT Trials 1-5. Thus, one would expect to find simultaneous increases in

performance on CVLT Trial 1 and CVLT Trials 1-5. There was also a positive correlation between the Beery VMI and the Beery VP. This is consistent with correlational data provided by the authors of these measures (Beery & Beery, 2006). Specifically, the authors demonstrated positive and significant correlations between the Beery VMI and the Beery VP for participants between the ages of 1 and 18 years, providing evidence for their part-whole intercorrelations hypothesis, which suggests that the Beery VP measures, to some extent, a part of what the Beery VMI measures (Beery & Beery, 2006).

Contrary to previous findings, positive correlations between CVLT Learning Slope and GPT Dominant and between FAS and Beery VP are not currently discussed or supported in the literature. This is also the case for the negative correlational finding that emerged between Symbol Search and Beery VMI and between CVLT Trial 1 and GPT Dominant. In light of these findings, it is important to highlight that bivariate correlations, for purposes of this study, were conducted as a preliminary means for determining the presence of multicollinear variables that did not provide a unique contribution to subsequent analyses. It is also important to note the risks inherent in attempting to interpret relationships culled from a small sample of participants (Baguley, 2012; Howell, 2007). Given the dearth of information regarding potential relationships between these variables in the literature, the primary purpose of the analysis, and the risks inherent in using small sample sizes, it is not possible to produce sound and practical interpretations for these findings.

With regard to relationships between task performance and demographics, participants who were older at the time of diagnosis were found to have obtained higher scores on the TMT A at follow-up when compared to baseline. From a developmental standpoint, one might expect for patients who were older at time of diagnosis to fare better in the long run on tasks that require some degree of attention given the caudal-to-rostral course of frontal lobe development and myelination (Dietrich et al., 1988). More specifically, the frontal lobe structures in these patients would have had more time to mature and myelinate, decreasing their vulnerability to insult. The data also revealed that females had a significantly higher decline in Digit Span than males. Several studies (see Lynn & Irwing, 2008) examining the relationship between sex and Digit Span performance have reported a slight advantage in females over males. This advantage was more pronounced in females who were 5 years of age than those between the ages of 6 and 16 years. Based on these studies, one would expect females to perform better than males at both baseline and follow-up. Analysis of the data, however, suggests the opposite. That is, males outperformed females at both time points. Moreover, while male performance declined by a marginal amount, female performance decrease by two-thirds of one standard deviation.

One possible explanation for the significant differential in performance on Digit Span is regression toward the mean, though an analysis of the data did not reveal extreme female performance scores at baseline. It is plausible, however, that the ratio of females to males and the presence of outliers impacted the results of the data. More specifically, a

higher proportion of females allowed for more variability in performance scores. Additionally, because there were a limited number of males (n = 8) in the study, any extreme changes in performance would have resulted in a greater impact on the mean. In this case, one male participant exhibited an increase of five scaled points, which is well over 1.5 standard deviations. Excluding this male's score would have resulted in a less drastic difference between males and females on this measure. While this participant's score did not fully account for the observed difference between the sexes on Digit Span, it seemed to play a contributory role, particularly when coupled with the notion of decreased variability resulting from unequal distributions.

Although females experienced higher drops in Digit Span, they were noted to have more significant increases on the GPT Non-Dominant task in comparison to males. Various studies (see Bryden & Roy, 2005) have indicated that females tend to perform better on GPT than males, with more pronounced differences observed for the non-preferred hand. It has been argued that differences in performance are rooted in the ability of females to manipulate pegs more easily than males given their relatively smaller finger size (see Bryden & Roy, 2005). Presumably, the advantage would be preserved in females with MS, given that finger size is not impacted by the disease. While plausible explanations for the bivariate relationships between task performance and demographic variables have been presented, the aforementioned caveat regarding the extrapolation of relationships based on small samples should be considered.

Neuropsychological Functioning

Individual-level performance. Examination of individual cognitive trajectories using the RCI method and frequency of cognitive impairment using the BPD method revealed that the majority of participants (55%) in the study exhibited more improvements than declines across the 16 neuropsychological measures. At follow-up, cognitive impairment, defined as impaired performance on three or more measures was observed in 9 of 20 participants (45%), a finding that is expected to occur by chance less than 5% of the time. It is important to note, however, that 45% of the sample also exhibited cognitive impairment at baseline. Interestingly, one participant who had demonstrated impairment on three tasks at baseline was impaired on only one task at follow-up. Conversely, another who had demonstrated impairment on only one task at baseline became impaired on three tasks upon reevaluation. When evaluating cognitive impairment data, it is important to consider the time between disease onset and baseline testing. In this study, for instance, many participants had been diagnosed several months or years prior to undergoing baseline testing. It is likely that by the time they were tested, participants had already experienced some cognitive changes as a results of the disease process (e.g., exacerbations), which might explain why 45% of participants were found to have significant cognitive impairment at baseline. It is also important to consider the time between baseline and follow-up. It is possible that 14 months is not a sufficient time frame to capture changes related to cognitive impairment, particularly when changes are expected to meet certain criteria, such as those set my Ingraham and Aiken (1996). Given

what we know about the disease process, it is likely that the percent of participants exhibiting cognitive impairment will increase over time.

With that said, the extent of cognitive impairment observed in the current study is comparable to the 41% frequency of impairment reported by MacAllister, Christodoulou and colleagues (2007). However, it is significantly lower than the 70% rate of impairment reported by Amato and colleagues (2010). Although the methods used to detect cognitive impairment between the present study and the Amato et al. (2010) study were relatively analogous, discrepancies in outcomes could reflect differences in the composition of the test batteries. Of the 15 measures in the Italian battery, only three (i.e., SDMT, TMT A, TMT B) were noted to overlap with those used in the current study. Interestingly, a large proportion of participants were noted to fail many of the non-overlapping tests (i.e., Semantic Verbal Fluency Tests, Token Test, Selective Reminding Test). Taken together, it is plausible that Amato et al. (2010) used measures with higher levels of sensitivity to MS-related pathology. Alternatively, cultural factors may have contributed to observed discrepancies. The level of impairment on measures of language, for instance, is noteworthy and potentially indicative of differences in levels of proficiency with the Italian language. While it is presumed that participants in the Amato et al. (2010) study were fluent Italian speakers, it is possible that the study enlisted children who spoke two languages (e.g., Italian/English). The differential performance between bilinguals and monolinguals on tasks of verbal fluency and vocabulary has certainly been evidenced in the literature (see Portocarrero, Burright, & Donovick, 2007). However, without having

any indication of the levels of proficiency of the participants in the Italian study, it is difficult to determine the extent to which differences in language might have played a role.

With regard to patterns of performance across neurocognitive domains, sensitivity to declines in working memory (Digit Span), visual scanning and sustained attention (SDMT), immediate verbal learning (CVLT Trial 1), and visual perception (Beery VP), were most pronounced (declines in ≥ 50% of the sample). Declines in shifting attention (TMT B) and motor coordination and visual organization (Beery VMI) were also evidenced in 40 to 45% of participants, respectively. These findings are generally consistent with prior longitudinal studies demonstrating prominent levels of deterioration in performance across the SDMT, the Beery VMI, the TMT B, and measures of verbal learning and memory (i.e., SRT) that are comparable to the CVLT. These findings are also partially supportive of the primary hypothesis that patients in the study would demonstrate declines in verbal fluency, complex attention (i.e., sustained, shifting), and verbal learning and memory.

Although a significant number of participants were noted to exhibit declines in working memory and visual perception, it is difficult to establish patterns of performance in these domains as it relates to pediatric MS patients since measures tapping these abilities have not been included in prior longitudinal studies. Interestingly, performance on Digit Span was noted as having one of the highest levels of impairment in a cross-sectional study (Julian et al., 2013) that included 25 measures. Thus, it is likely that

future studies will reveal rates of impairment in working memory that are comparable to those found in the current study. Another noteworthy finding is the relative preservation of performance on a task of verbal fluency. Although the majority of participants were noted to remain stable or improve on the FAS (70% combined), it is important to highlight that 30% of participants were still noted to exhibit declines. Thus, while the overall extent of improvement on this measure is inconsistent with findings from prior studies (i.e., Amato et al., 2010; MacAllister, Christodoulou, et al. 2007), the rate of decline on this task is comparable to the rate of participants who were categorized as impaired at follow-up (25%) in the MacAllister, Christodoulou, et al. (2007) study.

Group-level performance. Significant changes in neurocognitive functioning were also detected at the group level, although to a lesser extent than was observed at the individual level. As a whole, participants performed better on Digit Span at baseline when compared to follow-up, suggesting that pediatric MS patients exhibit significant cognitive declines in auditory working memory within a relatively short period of time. This finding is consistent with the rate of participants (65%) who were identified through the RCI method as having declined on this measure. On the contrary, group performance on TMT A was noted to be higher at follow-up in comparison to baseline, suggesting that pediatric MS patients exhibit improvements in visual scanning and motor speed over time. This finding is also consistent with the rate of participants (45%) who were identified through the RCI method as having improved on this measure.

Overall, results do not support the primary hypothesis and are only partially consistent with existing longitudinal studies suggesting more widespread deficits. Again, it is important to note that measures of working memory have not been included in prior longitudinal studies, making it difficult to determine the status and prevalence of working memory deficits in pediatric MS patients over time. Discrepancies between the present investigation and prior longitudinal studies may be attributable to differences in test-retest intervals (14.65 months in the present study compared with 21 to 24 months in prior studies). It is possible that deficits in areas other than working memory will emerge once participants are re-assessed a third time (i.e., 2 years post initial evaluation). It is also important to note that prior studies have primarily utilized frequency-based and individual-level methodologies as opposed to multivariate group-based analyses to detect neurocognitive declines over time. This is important as methods for evaluating cognitive change over time have been noted to yield measurably different outcomes (Walker et al., 2011).

Predictors of performance. Predictive relationships between all clinical variables of interest and several of the measures in the study were evidenced. As expected, higher reported levels of fatigue at baseline were noted to predict some degree of decline in functioning over time. Of the 16 measures, the TMT A emerged as most vulnerable to change as a function of fatigue. This can be expected given the degree of sustainability that is required on this task. It is interesting, however, that a similar relationships did not emerge between fatigue and other similarly demanding tasks,

considering the increase in fatigue-related symptomology over time, as evidenced by mean declines on the PedsQL MFS. Like fatigue, depression emerged as a significant predictor of declines in performance. The inverse relationship between depression and performance was most evidenced on tasks of visual perception (Beery VP) and manual speed and dexterity (GPT Dominant). While the impact of depression on cognitive functioning, particularly when left untreated, is well-established in the literature (Venkateswaran et al., 2011), the direct effects of depression on the aforementioned domains have not been identified, making it difficult to explain the emergence of these specific relationships. With that said, individuals with higher levels of depression have been noted to perform worse on less cognitively engaging or structured tasks that provide more opportunities to engage in negative, task-irrelevant thoughts (Kircanski, Joormann, & Gotlib, 2012). Although the Beery VP and GPT are relatively structured, it could be argued that they require relatively lower levels of attentional resources in comparison to tasks like Digit Span, TMT B, and SDMT, for example. Consequently, participants might have had more opportunities to become disengaged and detached from the tasks.

In addition to these expected outcomes, several relationships emerged that were contrary to what had been originally hypothesized. Higher levels of depression, for instance, were also noted to predict performance on tasks of immediate verbal learning (CVLT 1) and working memory (Digit Span). Interestingly, depression emerged as a predictor of positive, as opposed to negative, change over time. In addition to the already mentioned hypothesis related to differential engagement across tasks, it is possible that

participants who sought treatment for depressive symptoms exhibited sufficient improvements as to change the nature of the relationships between depression and performance. Although the multidisciplinary clinic from which participants were culled provide recommendations for treatment of mood-related symptoms, information regarding the use of antidepressants and/or therapeutic intervention was not collected as part of this study, making it difficult to verify this tentative supposition.

Like depression, age at onset, relapse rate, and disease duration were all noted to have unexpected relationships with several of the tasks. Participants who were older at time of diagnosis, for instance, were reported as being more likely to experience decreases in performance on a task of visual scanning and sustained attention (SDMT). This finding was unexpected considering what is known about the caudal-to-rostral course of frontal lobe development (Dietrich et al., 1988) and the presence of immature pathways and developing networks in younger patients that are more susceptible to MS-related disruptions (Fields, 2008; Julian et al., 2011). The finding that higher rates of relapse were predictive of improvements on a task of immediate verbal learning (CVLT Trial 1) was equally unexpected and contrary to what is presumed to occur after a patient has experienced multiple periods of inflammation, prolonged immunological activation, and demyelination (Compston & Coles, 2002; MacAllister et al., 2012).

Lastly, the present study found that participants who had been diagnosed with RRMS for longer periods of time performed better on tasks of immediate verbal learning (CVLT Trial 5 and CVLT B) over time. Existent data related to disease duration,

however, suggest the presence of more pronounced deficits in these patients resulting from more frequent exposure to periods of inflammation and demyelination, which are inevitably associated with the length of time since diagnosis (Banwell & Anderson, 2005; Julian et al., 2011; MacAllister et al., 2005). It is possible, however, that these patients had more time to psychologically adjust to the disease process, and that their favorable adjustment outweighed the effects of the process itself on these measures. With that said, all of these unexpected findings more likely represent spurious results caused by the implementation of complex analyses involving a small sample size and a relatively high number of variables. Violating Occam's razor by employing more complex statistical analyses in favor of those characterized by parsimony and fewer assumptions might have led to such findings (Baguley, 2012).

Implications

Findings from the present investigation converge, to some extent, with prior studies suggesting significant changes in the cognitive functioning of pediatric MS patients, with fatigue and depression as indicators of increased impairment over time. At a global level, these findings contribute to what is currently known about the impact of pediatric MS on cognitive performance over time. Specifically, this study enhances the literature base from which to extrapolate individual and group-based patterns of performance. Findings further elucidate the challenges inherent in studying an ongoing and active brain-based disease that transpires with and without obvious clinical

symptoms (e.g., exacerbations). More salient, however, are the real-world implications for pediatric MS patients, their families, and the professionals who serve them.

Results from this study, for instance, highlight the importance of routine (i.e., annual) surveillance of pediatric MS patients within the context of a multidisciplinary environment. Within this context, neuropsychologists should strive to detect and analyze the impact of psychological distress (i.e., depression) and fatigue on behavioral, academic, and social functioning. Neuropsychological screenings should also be conducted for purposes of monitoring and informing educational and treatment planning. Because health care reforms have given rise to the reduction of medical procedures and services that are perceived as being unnecessary, it will be incumbent upon business administrators, physicians, and other professionals who interact with insurance companies to advocate on behalf of patients in need of regular surveillance.

The extent of individual levels of fatigue, depression, and cognitive impairment observed in the current sample supports the need for intervention across settings. The implementations of school-based interventions are particularly important given the extent of time children spend in that setting. To facilitate the appropriate care of pediatric MS patients, members of the medical team should seek to consult with school personnel in order to provide information with regard to the disease process and the specific medical needs of the child. Because the course of MS is largely unpredictable and characterized by varying degrees of active clinical symptoms (Im-Wang, Milazzo, & Mowry, 2011), discussions around potential accommodations may be necessary.

Though professionals may wish to implement accommodations that address common difficulties experienced by children with MS, results from the current study support the use of individualized interventions to address the fluctuating needs of each child. Some children in the sample, for instance, were noted to experience declines in some areas that other children improved upon. Other children were noted to experience deficits in a specific area at baseline, but were noted to approximate age expectations at follow-up, and vice versa. Within and between differences in performance are expected given the waxing and waning course of MS, the range of medications prescribed to patients, the varying degrees of psychosocial support, differences in experience of fatigue, and other factors that contribute to the cognitive burden of MS. Interventions should be monitored and modified according to the specific, and potentially fluctuating, needs of each child in the course of his/her disease.

Lastly, it is important to note that while there was some degree of overlap between those areas noted as being most impaired in prior studies, and those identified in this study, some differences did emerge. Although pediatric batteries tend to emphasize areas most commonly purported to be affected in adults with MS (i.e., working memory, processing speed, and visual-spatial processing; Yeh et al., 2011) differences in findings support the inclusion of measures that assess a wide-range of domains, such as learning and memory, language, attention, executive functioning, and sensorimotor processes. Results also indicate the need for inclusion of measures that tap into a child's psychological functioning.

Limitations and Considerations

Although the present study was designed to provide novel and valuable information with regard to the nature of cognitive outcomes in pediatric MS over time, there are a number of limitations that should be highlighted. The first, and perhaps most significant limitation of the present study, is the modest sample size and potentially low statistical power. Although the sample size of the present study is comparable, if not greater, than many pediatric MS cohort studies in the literature, statistical power may not have been sufficient to detect all significant findings. Statistical power refers to the probability of rejecting the null hypothesis when the alterative hypothesis is true (Stevens, 2009). Small samples impact statistical power by decreasing the likelihood of finding accurate relationships between variables of interest. Although power was increased by utilizing a more liberal alpha level (i.e., .10), the limitations posed by small samples on results culled from complex statistical analyses cannot be underestimated (Mertler & Vannatta, 2005; Stevens, 2009). This caveat is particularly relevant to multiple regression analyses, wherein a ratio of 15 participants for every predictor variable is recommended (Mertler & Vannatta, 2005). According to this ratio recommendation, the present study would have needed to include 75 participants in order to obtain valid prediction results. Thus, plausible explanations for findings in the current study should be regarded as tentative hypotheses rather than conclusive generalizations about the pediatric MS population.

In addition to size, the composition of the sample poses threats to the generalizability or external validity of the study. Generalizability refers to the extent to which a study's results apply to those within and across populations (Heppner et al., 2008). Because of its limited characteristics, a homogenous sample decreases the extent of a study's generalizability. The current study includes a sample that is limited by several key characteristics including age, diagnosis, and geographic composition. Specifically, because the sample is comprised of pediatric MS patients, the extent to which results can generalize to the adult MS population as well as other clinical and non-clinical populations is limited. The data were also culled from patients who reside primarily in the southwestern region of the U.S. Thus, results may not generalize to pediatric MS patients in other geographic regions. As previously mentioned, adult research suggests that MS is a geographically-related disease with a latitudinal gradient (Ruggieri et al., 2011; Yeh et al., 2009). Thus, future studies should collect a more thorough demographic history that includes prior states and countries of residence; having this information will assist in determining whether the geographic distribution that is evidenced in adult studies is relevant to the MS pediatric population.

The repeated measures design of the study also poses some limitations. Specifically, studies that utilize repeated assessments are often plagued by two factors, namely, statistical regression toward the mean and practice effects (MacAllister, Christodoulou, et al., 2007). Statistical regression toward the mean is a phenomenon in which extreme variables tend to fall closer to the average range upon re-assessment;

whereas, practice effects refers to a phenomenon in which individual performance increases as a result of prior exposure to a task (Coaley, 2010). The compensatory nature of these factors might suppress actual deterioration in cognitive functioning to some degree. In other words, because true losses in cognitive functioning might be compensated by the gains made from these factors, actual impairments could be greater than scores would suggest (MacAllister, Christodoulou, et al., 2007). To ameliorate these effects, future studies could incorporate a healthy control comparison group, alternate test forms, and longer test-retest intervals.

With regard to alternate test forms, a recent meta-analytic study found significant reductions in test score inflations when alternate forms were utilized. With that said, it is important to note that alternate forms may not eliminate practice effects associated with learned strategies (Beglinger et al., 2005). Employing longer test-retest intervals could also be useful in distinguishing between explanations based on methodological artifacts versus true cognitive deterioration. However, using longer retest periods may also increase the extent to which findings are attributable to extraneous variables (e.g., treatment effects, family stressors) rather than true changes in functioning. Nevertheless, while the variability of practice effects reported in the literature have been largely dependent upon the domain of interest (i.e., higher practice effects evidenced on tests of learning and memory), a 12-month interval has generally been recommended for managing practice effects in adults and children (Salthouse & Tucker-Drob, 2008; Slade et al., 2008). With that said, several studies have found practice effects to persist 2 to 3

years (Calamia, Markon, & Tranel, 2012), 5 years (Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2008), or more than 7 years (Salthouse, Schroeder, & Ferrer, 2004) after testing. While these strategies were not implemented in the current study due to its archival nature, it is important to note that the recommended 12-month interval was observed.

Another important issue to highlight is that of psychometrics. The brief neuropsychological battery used in the present study encompasses a comprehensive assortment of tests designed to assess several neurocognitive processes with purported susceptibility to MS. With that said, a couple of the measures (i.e., TMT, GPT) are based on poor, incomplete, or dated standardization methods and/or psychometric properties (i.e., validity, reliability). One of the most critical aspects of standardization is the manner in which normative data is culled (Coaley, 2010). Normative data allow researchers to compare an individual's performance score with scores from a reference sample that is thought to represent the population of interest. When normative data are poorly populated or dated, it is difficult to identify and quantify differences that exist between individuals, over time, or across situations (Coaley, 2010). In addition to the presence of normative data, reliability and validity are two principal criteria utilized to evaluate the quality of quantitative measures (Lodico et al., 2006). Reliability refers to an instrument's ability to produce consistent and stable scores over time, administrations, and/or raters; whereas, validity refers to the instrument's ability to accurately measure the construct it intends or purports to assess. When a measure lacks in one of these properties, its consistency,

accuracy, and overall quality are called into question (Coaley, 2010). Despite this limitation, all performance-based cognitive measures from the brief battery were analyzed in the study. Additionally, the study analyzed the depression and fatigue inventories. However, the AI, which had significant psychometric issues, was only used for descriptive purposes. As the field of neuropsychology progresses, it will be important for test developers to construct measures with more rigorous psychometric properties and collect stratified normative data that accounts for shifts in cultural norms and aggregate changes in cognitive performance.

Finally, there are various factors that should be considered when interpreting data related to changes in cognitive functioning that were not controlled for in this study. Factors such as resiliency, treatment effects, and cognitive reserve are rarely discussed in studies related to the cognitive trajectory of MS, yet their potential impact cannot be overstated. Although the effects of stress on human functioning are well-established in the literature, (Yerkes & Dodson, 1908), studies specific to the impact of stress on cognitive functions such as attention, memory, and decision making (e.g., Broder, 2003; Chajut & Algom, 2003; Dougherty & Hunter, 2003; Larsen & Baddeley, 2003) have become more prevalent. These studies have also identified a number of moderating variables, protective factors, and behavioral strategies that appear to promote resilience to stress. Internal and external resilience factors have been implicated in performance outcomes of children with a wide-range of chronic conditions (e.g., Armstrong, Gyato, Awadalla, Lustig, & Tochner, 2004; McNally et al., 2013). Specifically, children with

optimistic dispositions, strong perceptions of self-efficacy, healthy coping strategies, and access to family/community supports have been noted to compensate for the severity and intensity of cognitive deficits secondary to their illness and its associated treatments more readily than children without these characteristics and supports (Armstrong et al., 2004; McNally et al., 2013). Taken together, differences in cognitive performance across individuals and cohorts may be partially attributable to differing levels of resilience.

Individual changes in cognitive performance could also be attributable to disease modifying treatment (DMT) regimens that may include first-line methods (i.e., glatiramer acetate, interferon beta-1a IM, interferon beta-1a SC, interferon beta-1b SC), as well as second-line or breakthrough therapies (e.g., natalizumab, mitoxantrone, cyclophosphamide, methotrexate; Pohl & Waubant, 2011; Yeh & Rodriguez, 2011). These therapies are generally used on the basis of their ability to rapidly inhibit the adverse effects of inflammation and reduce the accumulation of brain lesions, which positively impact the pathophysiology of cognitive dysfunction (Comi, 2010). Adult MS studies assessing the effects of glatiramer acetate, natalizumab, and the three classes of interferon on cognitive functioning (see Comi, 2010), have, in fact, demonstrated overall positive effects on cognitive alterations. With that said, effects on cognitive performance appear to be associated, to some extent, with dosing. In a study by Patti et al. (2010), for example, higher doses of interferon beta-1a SC were shown to predict lower cognitive impairment. It is also important to note that direct comparative studies between these therapies have not been performed in pediatric MS patients. Thus, it is not possible to

determine whether, and to what extent, one therapy option impacts cognitive performance relative to another. Lastly, very few data exist regarding the use of second-line therapies in pediatric MS. Thus, the contribution (positive or negative) of these therapies on the pathophysiology of cognitive dysfunction is largely unknown.

In addition to treatment effects, it is important to consider differences in cognitive reserve capacity. The theory of cognitive reserve maintains that individuals with higher capacity possess greater protection from cognitive impairment than those with lower capacity, given comparable brain insults (Palmer et al., 2003; Satz, 1993). Capacity is believed to vary as a function of brain size, quantity of synapses and dendritic branches, and neuronal efficiency. General intelligence is considered to be an indirect measure of capacity, in such that functioning is more likely to be preserved in individuals with higher pre-morbid IQ scores (Palmer et al., 2003; Satz, 1993). Considering this theory, it is plausible that participants in the study were less affected than individuals from prior studies as a result of differences in pre-morbid levels of intelligence. This, however, cannot be confirmed for two reasons. First, a measure of IQ was not utilized in the present study. Moreover, it is difficult to obtain accurate levels of pre-morbid intelligence as it is common for patients to be evaluated after, rather than prior to, disease onset.

Future Directions

Despite the aforementioned limitations, the current study provides important and enlightening findings that contribute to the fields of neuropsychology and pediatric multiple sclerosis. Of particular import to future studies will be the utilization of larger

cohorts to increase the statistical power and generalizability of findings. Given the rarity of pediatric MS, researchers should consider building partnerships with other study sites to increase the pool of participants from which to conduct studies. This endeavor, of course, has its unique set of obstacles, one of which would involve the utilization of a standardized battery across sites. Although the time needed to create and implement a standardized battery would inevitably lead to delays in the dissemination of longitudinal data, the literature with regard to cognitive changes in pediatric MS patients would be improved in the long-run.

Future studies should also include control groups comprised of age-, gender-, and SES-matched participants. Comparing pediatric MS patients against groups comprised of healthy participants could allow for a more accurate depiction of the effects of MS over time. Of additional concern is the dearth of literature regarding a standard set of criteria for determining whether a participant has reached the threshold of cognitive impairment. This issue has not only left researchers to make subjective decisions based on what they believe to be clinically meaningful, but has resulted in variable findings that are difficult to compare and evaluate. Consequently, it is imperative that the field maintain a consensus with regard to what constitutes cognitive impairment and what method researchers should employ to obtain criteria for determining classifications of patients. Operationalizing this term and its criteria will make it more feasible to directly compare performance outcomes across studies.

Studies aimed at assessing changes across the lifespan are also needed to more definitively address the long-term prognosis of pediatric MS patients. Longitudinal studies with multiple follow-up periods have the potential for assessing the emergence of deficits in occupational productivity and social functioning as well as establishing more accurate patterns of performance associated with specific cognitive domains. Investigations examining the relationship between DMTs, as well as other medications used for symptom management (e.g., anti-depressants), and cognitive performance in pediatric MS patients should also be conducted to further advance understanding of treatment as a contributory factor. Finally, studies aimed at elucidating the relationship between neurocognitive decline and neuroimaging indicators of disease progression (i.e., loss of brain volume, white and grey matter atrophy) are warranted.

Summary

The findings of this investigation are generally consistent with the conclusions of prior longitudinal studies regarding the presence of declines in cognitive performance and the relevance of clinical factors in predicting cognitive outcomes in pediatric MS patients. Findings, however, also reflect the variable nature of MS-related symptomology and outcomes, as some results were found to deviate from prior findings. Although not all results were found to have practical implications, the unique contributions of this study cannot be understated. This study, for example, was the first of its kind to identify significant declines in working memory at the individual and group levels, thus enhancing what is currently known about the trajectory and evolution of cognitive

functioning in pediatric MS patients. Findings culled from regression analyses also provided evidence for the adverse impact of fatigue and depression on cognitive performance over time. Moreover, through use of RCI, this investigation was able to identify whether observed declines in performance were meaningful and reliable. Lastly, although many studies have provided estimates of cognitive impairment, this investigation was the first pediatric MS study to describe and employ an empirical methodology by which to accomplish this goal. Taken together, these findings provide strong support for regular (i.e., annual) neuropsychological screenings and multidisciplinary care to address the unique medical, academic, and neuropsychological needs of each pediatric MS patient.

REFERENCES

Aaen, G. S. (2011). Corticosteroids in multiple sclerosis relapses. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 92-95). Cambridge, NY: Cambridge University Press.

Achenbach, T. M., & Rescorla, L. A. (2000). *Manual for the ASEBA preschool forms and profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.

Adams, W., & Sheslow, D. (1995). *Wide Range Assessment of Visual Motor Abilities*. Wilmington, DE: Wide Range, Inc.

Al-Futaisi, A. (2007). Childhood multiple sclerosis and related disorders. *Oman Medical Journal, 22*, 11-15.

Amato, M. P., Gorretti, B., Ghezzi, A., Lori, S., Zipoli, V., Moiola, L.,… Trojano, M. (2010). Cognitive and psychosocial features of childhood and juvenile MS: Two-year follow-up. *Neurology, 75*, 1134-1140. doi:10.1212/WNL.0b013e3181f4d821

Amato, M. P., Gorretti, B., Ghezzi, A., Lori, S., Zipoli, V., Portaccio, E.,… Lopez, M. (2008). Cognitive and psychosocial features of childhood and juvenile MS. *Neurology, 70*, 1891-1897.

Amato**,** M. P.**,** Ponziani**,** G.**,** Siracusa**,** G., & Sorbi, S. (2001). Cognitive dysfunction in early-onset multiple sclerosis: a reappraisal after 10 years. *Archives of Neurology, 58*, 1602-1607.

Amato, M. P., Zipoli, V., & Portaccio, E. (2008). Cognitive changes in multiple sclerosis. *Expert Review of Neurotherapeutics, 8*, 1585-1596. doi:10.1586/14737175.8.10.1585

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: Text revision* (4th ed.). Arlington, VA: American Psychiatric Association.

Anastasi, A. (1988). *Psychological testing* (6th ed.). New York, NY: MacMillan.

Army Individual Test Battery. (1944). *Manual of directions and scoring*. Washington, DC: War Department, Adjutant General's Office.

Armitage, S. G. (1946). An analysis of certain psychological tests used for the evaluation of brain injury. *Psychological Monographs, 60,* 30–34.

Armstrong, C. L., Gyato, K., Awadalla, A. W., Lustig, R., & Tochner, Z. A. (2004). A critical review of the clinical effects of therapeutic irradiation damage to the brain: The roots of controversy. *Neuropsychology Review, 14*, 65-86. doi:10.1023/B:NERV.0000026649.68781.8e

Army Individual Test Battery. (1944). *Manual of directions and scoring*. Washington, DC: War Department, Adjutant General's Office.

Arrondo, G., Alegre, M., Sepulcre, J., Iriarte, J., Artieda, J., & Villoslada, P. (2009). Abnormalities in brain synchronization are correlated with cognitive impairment in multiple sclerosis. *Multiple Sclerosis, 15*, 509-516. doi:10.1177/1352458508101321

Ascherio, A., & Munger, K. L. (2010). Epstein-Barr virus infection and multiple sclerosis. *Current Neurology and Neuroscience Reports, 9*, 405-401. doi:10.1007/s11481-010-9201-3

Ascherio, A., Zhang, S. M., Hernan, M. A., Olek, M. J., Coplan, P. M., Brodovicz, K., & Walker, A. M. (2001). Hepatitis B vaccination and the risk of multiple sclerosis. *New England Journal of Medicine, 344*, 327-332.

Avanzini, M. A., Ricci, A., Scaramuzza, C., Semino, L., Pagella, F., Castellazzi, A. M.,… Marseglia, G. L. (2006). Deficiency in INFgamma producing cells in adenoids of children exposed to passive smoke. *International Journal of Immunopathology and Pharmacology, 19*, 609-616.

Baguley, T. (2012). *Serious stats: A guide to advanced statistics for the behavioral sciences*. New York, NY: Palgrave Macmillan.

Banwell, B., & Anderson, P. E. (2005). The cognitive burden of multiple sclerosis in children. *Neurology, 64*, 891-894.

Banwell, B., Bar-Or, A., Cheung, R., Kennedy, J., Krupp, L. B., Becker, D. J., & Dosch, H. M. (2008). Abnormal T-cell reactivities in childhood inflammatory demyelinating diseases and type 1 diabetes. *Annals of Neurology, 63*, 98-111. doi:10.1002/ana.21244

Banwell, B., Ghezzi, A., Bar-Or, A., Mikaeloff, Y., & Tardieu, M. (2007). Multiple sclerosis in children: Clinical diagnosis, therapeutic strategies, and future. *Lancet Neurology, 6*, 887-902.

Banwell, B., Kennedy, J., Sadovnick, D., Arnold, D. L., Magalhaes, S., Wambera, K.,… Bar-Or, A. (2009). Incidence of acquired demyelination of the CNS in Canadian children. *Neurology, 72*, 232-239. doi:10.1212/01.wnl.0000339482.84392.bd

Banwell, B., Krupp, L., Kennedy, J., Tellier, R., Tenembaum, S., Ness, J.,… Farrell, K. (2007). Clinical features and viral serologies in children with multiple sclerosis: A multinational observational study. *Lancet Neurology, 6*, 773-781. doi:10.1016/S1474-4422(07)70196-5

Baron, I. S. (2004). *Neuropsychological evaluation of the child*. New York, NY: Oxford University Press.

Bass, N. H. (1968). Pathogenesis of myelin lesions in experimental cyanide encephalopathy. *Neurology, 18*, 167-177.

Beery, K. E., & Beery, N. A. (2006). *The Beery-Buktenica* Developmental *Test of Visual-Motor Integration: Administration, scoring, and technical manual* (5th ed.). Minneapolis, MN: NCS Pearson, Inc.

Beglinger, L. J., Gaydos, B., Tangphao-Daniels, O., Duff, K., Kareken, D. A., Crawford, J.,… Siemers, E. (2005). Practice effects and the use of alternate forms in serial neuropsychological testing. *Archives of Clinical Neuropsychology, 20,* 517–529. doi:10.1016/j.acn.2004.12.003

Bender, L. A. (1938). *A visual motor gestalt test and its clinical use: Research monograph number 3*. New York, NY: American Orthopsychiatric Association.

Benedict, R. H. B, Cookfair, D., Gavett, R., Gunther, M., Munschaeuer, 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*, 549-558. doi:10.1017/S1355617706060723

Benedict, R. H. B., & Zivadinov, R. (2007). Reliability and validity of neuropsychological screening and assessment strategies in MS. *Journal of Neurology, 254*, II/22-II/25. doi:10.1007/s00415-007-2007-4

Benton, A. L., & Hamsher, K. (1976). *Multilingual Aphasia Examination.* Iowa City, IA: University of Iowa.

Benton, A. L., Hamsher, K. & Sivan, A. B. (1994). *Multilingual Aphasia Examination: Manual of instructions* (3rd ed.). Iowa City, IA: AJA Associates, Inc.

Boiko, A. N., Gusev, E. I., Sudomoina, M. A., Alekseenkov, A. D., Kulakova, O. G., Bikova, O.V.,… Favorova, O. O. (2002). Association and linkage of juvenile MS with HLA-DR2(15) in Russians. *Neurology, 58*, 658-660. doi:10.1212/WNL.58.4.658

Boiko, A., Vorobeychik, G., Paty, D., Devonshire, V., & Sadovnick, S. (2002). Early onset multiple sclerosis: A longitudinal study. *Neurology, 159*, 1006-1010.

Boringa, J. B., Lazeron, R. H. C., Reuling, I. E. W., Ader, H. J., Pfennigs, L., Lindeboom, J.,… Polman, C. H. (2001). The Brief Repeatable Battery of Neuropsychological Tests: Normative values allow application in multiple sclerosis clinical practice. *Multiple Sclerosis, 7*, 263-267.

Bornstein, R. A. (1985). Normative data on selected neuropsychological measures from a nonclinical sample. *Journal of Clinical Psychology, 41*, 651-659.

Boyd, J. R., & MacMillan, L. J. (2005). Experiences if children and adolescents living with multiple sclerosis. *The Journal of Neuroscience Nursing, 37*, 334-342.

Broder, A. (2003). Decision making with the "adaptive toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 611–625.

Bryden, P. J., & Roy, E. A. (2005). A new method of administering the Grooved Pegboard Test: Performance as a function of handedness and sex. *Brain and Cognition, 58*, 258-268.

Calamia, M., Markon, K., & Tranel, D. (2012). Scoring higher the second time around: Meta-analyses of practice effects in neuropsychological assessment. *The Clinical Neuropsychologist, 26*, 543-570. doi:10.1080/13854046.2012.680913

Callen, D. J., Shroff, M. M., Branson, H. M., Li, D. K., & Lotze, T., Stephens, D., & Banwell, B. L. (2009). Role of MRI in the differentiation of ADEM from MS in children. *Neurology, 72*, 968-973. doi:10.1212/01.wnl.0000338630.20412.45

Callen, D. J., Shroff, M. M., Branson, H. M., Lotze, T., Stephens, D., & Banwell, B. L. (2009). MRI in the diagnosis of pediatric multiple sclerosis. *Neurology, 72*, 961-967. doi:10.1212/01.wnl.0000338629.01627.54

Carswell R. (1838) Pathological anatomy: Illustrations of the elementary forms of the disease. London, England: Longman, Orme, Brown, Green and Longman.

Centofanti, C. C. (1975). *Selected somatosensory and cognitive test performances as a function of age and education in normal and neurologically abnormal adults* (Doctoral dissertation). Retrieved from http://ezproxy.twu.edu:2048/docview/302787501?accountid=7102

Chabas, D., Castillo-Trevino, T., Mowry, E. M., Strober, J. B., Glenn, O. A., & Waubant, E. (2008). Vanishing MS T2-bright lesions before puberty: A distinct MRI phenotype? *Neurology, 71*, 1090-1093.

Chabas, D., Krupp, L. B., & Tardieu, M. (2011). Controversies around the current operational defenitions of pediatric MS, ADEM, and related diseases. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 10-18). Cambridge, NY: Cambridge University Press.

Chajut, E., & Algom, D. (2003). Selective attention improves under stress: Implications for theories of social cognition. *Journal of Personality and Social Psychology, 85*, 231–248.

Charcot, J. M. (1877). Lectures on the disease of the nervous system delivered at La Salpetriere. London, England: New Sydenham Society.

Chelune, G. J. (2003). *Assessing reliable neuropsychological change*. In R. D. Franklin (Ed.), *Prediction in forensic and neuropsychology: Sound statistical practices* (pp. 123–147). Mahwah, NJ: Lawrence Erlbaum Associates.

Chelune G. J., Naugle R. I., Lüders, H., Sedlak, J., & Awad, I. A. (1993). Individual change after epilepsy surgery: practice effects and base-rate information. *Neuropsychology, 7,* 41-52.

Chitnis, T., Glanz, B., Jaffin, S., & Healy, B. (2009). Demographic of pediatric onset multiple sclerosis in an MS center population from the Northeastern United States. *Multiple Sclerosis, 15*, 627-631. doi:10.1177/1352458508101933

Chitnis, T., & Bar-Or, A. (2011). Pediatric MS: Biological presentation and research update. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 157-168). Cambridge, NY: Cambridge University Press.

Coaley, K. (2010). *An introduction to psychological assessment and psychometrics*. Thousand Oaks, CA: Sage Publications.

Cohen, M. (1997). *Children's Memory Scale*. San Antonio, TX: The Psychological Corporation.

Cohen, R. A., Paul, R., Zawacki, T. M., Moser, D. J., Sweet, L., & Wilkinson, H. (2001). Emotional and personality changes following cingulotomy. *Emotion, 1*, 83-50. doi:10.1037//1528-3542.1.j.38

Cole, G. F., Auchterlonie, L. A., & Best, P. V. (1995). Very early onset multiple sclerosis. *Developmental Medicine and Child Neurology, 37*, 667-672.

Colom, R., Burgaleta, M., Roman, F. J., Karama, S., Alvarez-Linera, J., Abad, F. J.,… Haier, R. J. (2013). Neuroanatomic overlap between intelligence and cognitive factors: Morphometry methods provide support for the key role of the frontal lobes. *NeuroImage, 72*, 143-152. doi: 10.1016/j.neuroimage.2013.01.032

Comi, G. (2010). Effects of disease modifying treatments on cognitive dysfunction in multiple sclerosis. *Neurological Sciences, 31*, 261-264. doi:10.1007/s10072-010-0436-9

Compston, A., & Coles, A. (2002). Multiple sclerosis. *Lancet, 359*, 1221-1231.

Compston, A., & Coles, A. (2008). Multiple sclerosis. *Lancet, 372*, 1502-1517.

Conners, C. K. (1997). *Conners' Rating Scales-Revised*. North Towanda, NY: Multi-Health Systems.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Fort Worth, TX: Harcourt Brace Jovanovich College.

Crockett, D. J., Hurwitz, T., & Vernon-Wilkinson (1990). Differences in neuropsychological performance in psychiatric, anterior- and posterior-cerebral dysfunctioning groups. *International Journal of Neuroscience, 52*, 45-57.

Dale, R. C., de Sousa, C., Chong, W. K., Cox, T. C., Harding, B., & Neville, B. G. (2000). Acute disseminated encephalomyelitis, multiphasic disseminated encephalomyelitis and multiple sclerosis in children. *Brain: A Journal of Neurology, 123*, 2407-2422.

Dean, G., Yeo, T. W., Goris, A., Taylor, C. J., Goodman, R. S., Elian, M.,… Compston, D. A. (2008). HLA-DRB1 and multiple sclerosis in Malta. *Neurology, 70*, 101-105.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis Kaplan Executive Function System: Technical manual*. San Antonio, TX: The Psychological Corporation.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test: Manual*. San Antonio, TX: The Psychological Corporation.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1994). *California Verbal Learning Test-Children's Version: Manual*. San Antonio, TX: The Psychological Corporation.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *California Verbal Learning Test-Second Edition: Manual*. San Antonio, TX: The Psychological Corporation.

des Rosiers, G., & Kavanaugh, D. (1987). Cognitive assessment in closed head injury: Stability, validity and parallel forms for two neuropsychological measures of recovery. *International Journal of Clinical Neuropsychology, 9*, 162-173.

DeStefano, F., Verstraeten, T., Jackson, L. A., Okoro, C. A., Benson, P., Black, S. B.,… Chen, R. T. (2003). Vaccinations and risk of central nervous system demyelinating diseases in adults. *Archives of Neuropsychology, 60*, 504-509.

DeStefano, F., Weintraub, E. S., & Chen, R. T. (2005). Recombinant hepatitis B vaccine and the risk of multiple sclerosis: A prospective study. *Neurology, 64*, 1317.

Dietrich, R. B., Bradley, W. G., Zaragoza, E. J., Otto, R. J., Taira, R. K., Wilson, G. H.,… Kangarloo, H. (1988). MR evaluation of early myelination patterns in normal and developmentally delayed infants. *American Journal of Roentgenology, 150*, 889-896.

Dikmen, S. S., Machamer, J. E., Winn, H. R., & Temkin, N. R. (1995). Neuropsychological outcomes at 1-year post head injury. *Neuropsychology, 9*, 80-90.

Di Sclafani, V., Mackay, R. D. S., Meyerhoff, D. J., Norman, D., Weiner, M. W., & Fein, G. (1997). Brain atrophy in HIV infection is more strongly associated with CDC clinical stages than with cognitive impairment. *Journal of the International Neuropsychological Society, 3*, 276-287.

Dougherty, M. R. P., & Hunter, J. (2003). Probability judgment and subadditivity: The role of working memory capacity and constraining retrieval. *Memory & Cognition, 31*, 968–982.

Dunn, L. M. (1959). *Peabody Picture Vocabulary Test*. Circle Pines, MN: American Guidance Service, Inc.

Dye, O. A. (1979). Effects of practice on Trail Making Test performance. *Perceptual and Motor Skills, 48*, 296.

Ehrenstein, W. H., Heister, G., & Cohen, R. (1982). Trail Making Test and visual search. *Archiv Für Psychiatrie Und Nervenkrankheiten, 231*, 333-338.

Fals-Stewart, W. (1992). An interrater reliability study of the Trail Making Test (Parts A and B). *Perceptual and Motor Skills, 74*, 39-42.

Feinstein, A. (2011). Multiple sclerosis and depression. *Multiple Sclerosis Journal, 17*, 1276–1281. doi:10.1177/1352458511417835

Ferreira, M. L., Machado, M. I., Dantas, M. J., Moreira, A. J., & Souza, A. M. (2008). Pediatric multiple sclerosis: Analysis of clinical and epidemiological aspects according to National MS Society Consensus. *Arquivos de Neuro-Psiquiatria, 66*, 665-670. doi:10.1590/S0004-282X2008000500011

Fields, R. D. (2008). White matter in learning, cognition, and psychiatric disorders. *Trends in Neuroscience, 31*, 361-370. doi:10.1016/j.tins.2008.04.001

Fielding, J., Kilpatrick, T., Millist, L., & White, O. (2009). Multiple sclerosis: Cognition and saccadic eye movements. *Journal of the Neurological Sciences, 277*, 32-36.

Findeis, M. K., & Weight, D. G. (1994). *Meta-norms for Indiana-Reitan Neuropsychological Test Battery and Halstead-Reitan Neuropsychological Test Battery for Children, ages 5-14*. Unpublished manuscript.

Flanagan, D. P., & Kaufman, A. S. (2004). *Essentials of the WISC-IV assessment*. Hoboken, NJ: John Wiley & Sons, Inc.

Francis, D. A., Bain, P., Swan, A. V., Hughes, R. A. C. (1991). An assessment of disability rating scales used in multiple sclerosis. *Archives of Neurology, 48*, 299-301.

French Research Group on Multiple Sclerosis (1992). Multiple sclerosis in 54 twinships: Concordance rate is independent of zygosity. *Annals of Neurology, 32*, 724-727.

Fromm-Auch, D., & Yeudall, L. T. (1983). Normative data for the Halstead-Reitan Neuropsychological Tests. *Journal of Clinical Neuropsychology, 5*, 221-238.

Gazzaniga, M. S., Ivry, R. B., Mangun, G. R., & Steven, M. S. (1998). *Cognitive neuroscience: The biology of the mind.* New York, NY: WW Norton.

Geurts, J. J., Bo, L., Roosendaal, S. D., Hazes, T., Daniels, R., Barkhof, F.,… van der Valk, P. (2007). Extensive hippocampal demyelination in multiple sclerosis. *Journal of Neuropsychology and Experimental Neurology, 66*, 819-827. doi:10.1097/nen.0b013e3181461f54

Ghassemi, R., Antel, S. B., Narayanan, S., Francis, S. J., Bar-Or, A., Sadovnick, A.D.,… Arnold, D. L. (2008). Lesion distribution in children with clinically isolated syndromes. *Annals of Neurology, 63*, 401-405. doi:10.1002/ana.21322

Ghezzi, A. (2004). Clinical characteristics of multiple sclerosis with early onset. *Journal of the Neurological Sciences, 25*, S336-S339. doi:10.1007/s10072-004-0336-y

Ghezzi, A., & Banwell, B. (2011). Clinical presentation of multiple sclerosis in children. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 36-47). Cambridge, NY: Cambridge University Press.

Goldstein G., Watson J. R. (1989). Test-retest reliability of the Halstead-Reitan Battery and the WAIS in a neuropsychiatric population. *The Clinical Neuropsychologist, 3*, 265-272.

Groff, M. G., & Hubble, L. M. (1981). A factor analytic investigation of the Trail Making Test. *Clinical Neuropsychology, 3*, 11-13.

Hahn, J. S., & Tenembaum, S. (2011). Clinical and biological features of acute disseminated encephalomyelitis. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 183-201). Cambridge, NY: Cambridge University Press.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall, Inc.

Halstead, W. C. (1947). *Brain and intelligence: A quantitative study of the frontal lobes.* Chicago, IL: University of Chicago Press.

Hambleton, R. K. (1993). Principles and selected applications of item response theory. In R. L. Linn (Ed.), *Educational measurement* (pp. 147-220). Phoenix, AZ: Oryx Press.

Hammill, D. D., Pearson, N. A., & Voress, J. K. (1993). *Developmental tests of visual perception* (2nd ed.). Austin, TX: Pro-Ed.

Harrison, P. L., & Oakland, T. (2000). *Adaptive Behavior System-Second Edition*. San Antonio, TX: The Psychological Corporation.

Harwood, R. H., Gompertz, P., & Ebrahim, S. (1994). Handicap one year after a stroke: Validity of a new scale. *Journal of Neurology, Neurosurgery, & Psychiatry, 57*, 825–829.

Hauser, S. L., Dawson, D. M., Lehrich, J. R., Beal. M. F., Kevy, S. V., Propper, R. D., … Weiner, H. L. (1983). Intensive immunosuppression in progressive multiple sclerosis: A randomized, three-arm study of high-dose intravenous cyclophosphamide, plasma exchange, and ACTH. *The New England Journal of Medicine, 308*, 173-180.

Healy, B. C., Ali, E. N., Guttmann, C. R., Chitnis, T., Glanz, B. I., Buckle, G.,… Ascherio, A. (2009). Smoking and disease progression in multiple sclerosis. *Archives of Neurology, 66*, 858-864.

Heaton, R. K. (1981). *Wisconsin Card Sorting Test manual.* Odessa, FL: Psychological Assessment Resources, Inc.

Heaton, R. K., Grant, I., & Matthews, C. G. (1991). *Comprehensive norms for an expanded Halstead-Reitan Battery: Demographic corrections, research findings, and clinical applications*. Odessa, FL: Psychological Assessment Resources.

Heaton, R. K., Temkin, N., Dikmen, S., Avitable, N., Taylor, M. J., Marcotte, T. D., & Grant, I. (2001). Detecting change: A comparison of three neuropsychological methods, using normal and clinical samples. *Archives of Clinical Neuropsychology, 16*, 75-91.

Heilbronner, R. L, Henry, G. K., Buck, P., Adams, R. L., & Fogle, T. (1991). Lateralized brain damage and performance on Trail Making A and B, Digit Span Forward and Backward, and TPT memory and location. *Archives of Clinical Neuropsychology, 6*, 251-258.

Hedstrom, A. K., Baarnhielm, M., Olsson, T., & Alfredsson, L. (2009). Tobacco smoking, but not Swedish snuff use, increases the risk of multiple sclerosis. *Neurology, 73*, 696-701. doi:10.1212/WNL.0b013e3181b59c40

Heppner, P. P., Wampold, B. E., & Kivlighan, D. M. (2008). *Research design in counseling* (3rd ed.). Belmont, CA: Thomson Brooks/Cole.

Holland, A. A., Graves, D., Greenberg, B. M., & Harder, L. L. (2012). Fatigue, emotional functioning, and executive dysfunction in pediatric multiple sclerosis. *Child Neuropsychology, 2012*, 1-15. doi:10.1080/09297049.2012.748888

Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 129-145). Hillsdale, NJ: Erlbaum.

Hollingshead, A. B. (1957). *Two factor index of social position*. New Haven, CT: Yale University.

Hsu, L. M. (1995). Regression toward the mean associated with measurement error and the identification of improvement and deterioration in psychotherapy. *Journal of Consulting and Clinical Psychology, 63*, 141–144.

Homack, S., Lee, D., & Riccio, C. A. (2005). Test review: Delis-Kaplan Executive Function System. *Journal of Clinical and Experimental Neuropsychology, 27*, 599-609.

Howell, D. C. (2007). *Statistical methods for psychology* (6th ed.). Belmont, CA: Thomson Wadsworth.

Im-Wang, S., & Milazzo, M. (2011). Living with pediatric multiple sclerosis: Patient well-being. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 144-156). Cambridge, NY: Cambridge University Press.

Ingraham, L. J., & Aiken, C. B. (1996). An empirical approach to determining criteria for abnormality in test batteries with multiple measures. *Neuropsychology, 10*, 120-124.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12-19.

Jaquiery, E., Jilek, S., Schluep, M., Meylan, P., Lysandropoulos, A., Pantaleo, G., & Du Pasquier, R. A. (2010). *European Journal of Immunology, 40,* 878-887.

Jellife, S. E., & White, W. A. (1919). *Disease of the nervous system: A textbook of neurology and psychiatry* (3rd ed.). Philadelphia, PA: Lea & Febiger.

Julian, L. J., & Arnett, P. A. (2009). Relationships among anxiety, depression, and executive functioning in multiple sclerosis. *The Clinical Neuropsychologist, 23*, 794-804. doi:10.1080/13854040802665808

Julian, L. J., & Mohr, D. C. (2006) Cognitive predictors of response to treatment for depression in multiple sclerosis. *The Journal of Neuropsychiatry and Clinical Neurosciences, 18*, 356-363. doi:10.1176/appi.neuropsych.18.3.356

Julian, L. J., Serafin, D., Charvet, L., Ackerson, J., Benedict, R., Braaten, E.,… Krupp, L. B. (2013). Cognitive impairment occurs in children and adolescents with multiple sclerosis: Results from a United States network. *Journal of Child Neurology, 28*, 102-107. doi:10.1177/0883073812464816

Julian, L. J., Trojano, M., Amato, M. P., & Krupp, L. B. (2011). Cognitive dysfunction in pediatric-onset multiple sclerosis. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 134-143). Cambridge, NY: Cambridge University Press.

Keith, R. A., Granger, C. V., Hamilton, B. B., & Sherwin F. S. (1987). The functional independence measure: A new tool for rehabilitation. In M. G. Eisenberg & R. C. Grzesiak (Eds.), *Advances in clinical rehabilitation* (pp. 6-18). New York, NY: Springer-Verlag.

Ketelslegers, I. A., Nuteboom, R. F., Boon, M., Castman-Berrevoets, C. E., Hintzen, R. Q., & Dutch Pediatric MS Study Group (2010). A comparison of MRI criteria for diagnosing pediatric ADEM and MS. *Neurology, 74*, 1412-1415. doi:10.1212/WNL.0b013e3181dc138b

Kircanski, K., Joormann, J., & Gotlib, I. H. (2012). Cognitve aspects of depression. *Wiley Interdisciplinary Reviews Cognitive Science, 3*, 301-313.

Kis, B., Rumberg, B., & Berlit, P. (2008). Clinical characteristics of patients with late-onset multiple sclerosis. *Journal of Neurology, 255*, 697-702.

Klove, H. (1963). Clinical neuropsychology. In F. M. Foster (Ed.), *The medical clinics of North America* (pp. 1647-1658). New York, NY: Saunders.

Knights, R. M., Ivan, L. P., Ventureyra, E. C. G., Bentivoglio, C., Stoddart, C., Winogron, W., & Bawden, H. N. (1991). The effects of head injury in children on neuropsychological and behavioral functioning. *Brain Injury, 5*, 339-351.

Knights, R. M., & Moule, A. D. (1968). Normative data on the motor steadiness battery for children. *Perceptual and Motor Skills, 26*, 643-650.

Krull, V. (1968) Consistent sex differences in a specific (decoding) test performance. Unpublished manuscript.

Krull, K. R., Okcu, M. F., Potter, B., Jain, N., Dreyer, Z., Kamdar, K., & Brouwers, P. (2008). Screening for neurocognitive impairments in pediatric cancer long-term survivors. *Journal of Clinical Oncology, 26*, 4138-4143. doi:10.1200/JCO.2008.16.8864

Krupp, L. B., & Elkins, L. E. (2000). Fatigue and declines in cognitive functioning in multiple sclerosis. *Neurology, 55*, 934-939.

Krupp, L. B., McLinskey, N., & Troell, R. (2008). Racial and ethnic findings in pediatric MS: An update. [Abstract]. *Neurology, 70*, A135.

Kujala, P., Portin, R., & Ruutiainen, J. (1997). The progress of cognitive decline in multiple sclerosis. *Brain, 120*, 289-297.

Kuntz, N. L., Chabas, D., Weinstock-Guttman, B., Chitnis, T., Yeh, A. E., Krupp, L.,… Waubant, E. (2010). Treatment of multiple sclerosis in children and adolescents. *Expert Opinion on Pharmacotherapy, 11*, 505-520. doi:10.1517/14656560903527218

Kurtzke, J. F. (1983). Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). Neurology, 33, 1444-1452.

Kutner, M.H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models* (5th ed.). New York, NY: McGraw-Hill.

Lafleche, G., & Albert, M. S. (1995). Executive function deficits in mild Alzheimer's disease. *Neuropsychology, 9*, 313-320.

Lai, J. S., Cella, D., Kupst, M. J., Holm, S., Kelly, M. E., Bode, R. K., & Goldman, S. (2007). Measuring fatigue for children with cancer: Development and validation of the pediatric functional assessment of chronic illness therapy-fatigue (pedsFACIT-F). *Journal of Pediatric Hematology/Oncology, 29*, 471-479.

Larsen, J. D., & Baddeley, A. (2003). Disruption of verbal STM by irrelevant speech, articulatory suppression, and manual tapping: Do they have a common source. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 56A,* 1249-1268.

Lewis, R., & Kupke, T. (1992) Intermanual differences on skilled and unskilled motor tasks in non-lateralized brain dysfunction. *The Clinical Neuropsychologist, 6*, 374–382.

Lezak, M. D. (1982). *The test-retest stability and reliability of some tests commonly used in neuropsychological assessment*. Paper presented at the 5th European Conference of the International Neuropsychological Society, Deauville, France.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York, NY: Oxford University Press.

Li, H., Rosenthal, R., & Rubin, D. B. (1996). Reliability of measurement in psychology: From Spearman-Brown to maximal reliability. *Psychological Methods, 1*, 98-107.

Lodico, M. G., Spaulding, D. T, & Voegtle, K. H. (2006). *Methods in educational research: From theory to practice*. San Francisco, CA: John Wiley & Sons, Inc.

Lunemann, J. D., Huppke, P., Roberts, S., Bruck, W., Gartner, J., & Munz, C. (2008). Broadened and elevated humoral immune response to EBNA1 in pediatric multiple sclerosis. *Neurology, 71*, 1033-1035. doi:10.1212/01.wnl.0000326576.91097.87

Lupien, S. J., & McEwan, B. S. (1997). The acute effects of corticosteroids on cognition: Integration of animal and human model studies. *Brain Research Reviews, 24*, 1-27. doi:10.106/S0165-0173(97)0004-0

Lynn, R., & Irwing, P. (2008). Sex differences in mental arithmetic, digit span, and g defined as working memory capacity. *Intelligence, 36*, 226-235. doi:10.1016/j.intell.2007.06.002

MacAllister, W. S., Belman, A. L., Milazzo, M., Weisbrot, D. M., Christodoulou, C., Scherl, T. E.,… Krupp, L. B. (2005). Cognitive functioning in children and adolescents with multiple sclerosis. *Neurology, 64*, 1422-1425. doi:10.1212/01.WNL.0000158474.24191.BC

MacAllister, W. S., Boyd, J. R., Holland, N. J., Milazzo, M.C., & Krupp, L. B. (2007). The psychosocial consequences of pediatric multiple sclerosis. *Neurology, 68*, S66-S69. doi: 10.1212/01.wnl.0000259420.54635.63

MacAllister, W. S., Christodoulou, C., Milazzo, M., & Krupp, L. B. (2007). Longitudinal neuropsychological assessment in pediatric multiple sclerosis. *Developmental Neuropsychology, 32*, 625-644.

MacAllister, W. S., Christodoulou, C., Milazzo, M., Preston, T. E., Serafin, D., Krupp, L. B., & Harder, L. (2012). Pediatric multiple sclerosis: What we know and where are we headed? *Child Neuropsychology, 2012*, 1-22. doi:10.1080/09297049.2011.639758

MacAllister, W. S., Christodoulou, C., Troxell, R., Milazzo, M., Block, P., Preston, T. E.,… Krupp, L. B. (2009) Fatigue and quality of life in pediatric multiple sclerosis. *Multiple Sclerosis, 15*, 1502-1508. doi:10.1177//1352458509345902

MacAllister, W. S., & Krupp, L. B. (2005). Multiple sclerosis-related fatigue. *Physical Medicine and Rehabilitation Clinics of North America, 16*, 483-502. doi:10.1016/j.pmr.2005.01.014

Mahoney F. I., & Barthel D. W. (1965). Functional evaluation: The Barthel Index. *Maryland State Medical Journal, 14*: 61–65.

Maj, M., D'Elia, L. F., & Satz, P. (1993). Evaluation of two new neuropsychological tests designed to minimize cultural bias in the assessment of HIV-1 seropositive persons: A WHO study. *Archives of Clinical Neuropsychology, 8*, 123-135.

Marrie, R. A., Yu, N., Blanchard, J., Leung, S., & Elliott, L. (2010). The rising prevalence and changing age distribution of multiple sclerosis in Manitoba. *Neurology, 74*, 465-471. doi:10.1212/WNL.0b013e3181cf6ec0

Marrosu, M. G., Sardu, C., Cocco, E., Costa, G., Murru, M. R., Mancosu, C.,… Contu, P.(2004). Bias in parental transmission of the HLA-DR3 allele in Sardinian multiple sclerosis. *Neurology, 63*, 1084-1086.

Matthews, C., Cleeland, C. S., & Hopper, C. L. (1970). Neurological patterns in multiple sclerosis. *Diseases of the Nervous System, 1970*, 161-170.

Matthews, C. G., & Klove, H. (1964). *Instruction manual for the Adult Neuropsychological Test Battery*. Madison, WI: University of Wisconsin Medical School.

McAdam, L. C., Blaser, S. I., & Banwell, B. L. (2002). Pediatric tumefactive demyelination: Case series and review of the literature. *Pediatric Neurology, 26*, 18-25. doi:10.1016/S0887-8994(01)00322-8

McCaffrey, R, J., Krahula, M. M., Heimberg, R. G., Keller, K. E., & Purcell, M. J. (1988). A comparison of the Trail Making Test, Symbol Digit Modalities Test, and the Hooper Visual Organization Test in an inpatient substance abuse population. *Archives of Clinical Neuropsychology, 3*, 181-187.

McCann, K., Farmer, J. F., & Patel, N. (2004). Childhood-onset multiple sclerosis and mood disorders: A case study. *Child Neuropsychology, 10*, 102-116.

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.

McKeon, A., & Lotze, T. (2011). Neuromyelitis optica in children. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 255-271). Cambridge, NY: Cambridge University Press.

McNally, K. A., Dietrich, A., Nuss, K., Wright, M., Taylor, H. G., Bangert, B, …Yeates, K. O. (2013). Injury versus noninjury factors as predictors of postconcussive symptoms following mild traumatic brain injury in children. *Neuropsychology, 27*, 1-12. doi:10.1037/a0031370

Mertler, C. A., & Vannatta, R. A. (2005). *Advanced and multivariate statistical methods: Practical application and interpretation* (3rd ed.). Glendale, CA: Pyrczak Publishing.

Mikaeloff, Y., Caridade, G., Assi, S., Suissa, S., & Tardieu, M. (2006). Prognostic factors for early severity in a childhood multiple sclerosis cohort. *Pediatrics, 118*, 1133-1139.

Mikaeloff, Y., Caridade, G., Rossier, M., Suissa, S., & Tardieu, M. (2007). Hepatitis B vaccination and the risk of childhood-onset multiple sclerosis. *Archives of Pediatric & Adolescent Medicine, 161*, 1176-1182.

Mikaeloff, Y., Caridade, G., Suissa, S., & Tardieu, M. (2009). Hepatitis B vaccine and the risk of CNS inflammatory demyelination in childhood. *Neurology, 72*, 873-880.

Mikaeloff, Y., Caridade, G., Tardieu, M., & Suissa, S, & KIDSEP study group. (2007). Parental smoking at home and the risk of childhood-onset multiple sclerosis in children. *Brain, 130*, 2589-2595.

Miller, D. C. (2010). *Best practices in school neuropsychology: Guidelines for effective practice, assessment, and evidence-based intervention*. Hoboken, NJ: John Wiley & Sons, Inc.

Mitrushina, M., Boone, K. B., Razan, J., & D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment* (2nd ed.). New York, NY: Oxford University Press.

Mumford, C. J., Wood, N. W., Kellar-Wood, H., Thorpe, J. W., Miller, D. H., & Compston, D. A. (1994). The British Isles survey of multiple sclerosis in twins. *Neurology, 44*, 11-15.

Murray, T. J. (2005). *Multiple sclerosis: The history of a disease*. New York, NY: Demos Medical Publishing.

Murray, T. J. (2009). The history of multiple sclerosis: The changing frame of the disease over centuries. *Journal of Neurological Sciences, 277*, S3-S8. doi:10.1016/S0022-510X(09)70003-6

Nagy, Z., Westerberg, H., & Klingberg, T. (2004). Maturation of white matter is associated with the development of cognitive functions during childhood. *Journal of Cognitive Neuroscience, 16*, 1227-1233.

Nakai, M., & Ke, W. (2009). Statistical models of longitudinal data analysis. *Applied Mathematical Sciences, 3*, 1979-1989.

Newcomer, J. W., Craft, S., Hershey, T., Askins, K., & Bardgett, M. E. (1994). Glucocorticoid-induced impairment in declarative memory performance in adult humans. *The Journal of Neuroscience, 14*, 2047-2053.

Neuteboom, R. F., Boon, M., Castman-Berrevoets, C. E., Vles, J. S., Gooskens, R. H., Stroink, H.,… Hintzen, R. (2008). Prognostic factors after a first attack of inflammatory CNS demyelination in children. *Neurology, 71*, 967-973.

O'Connor, K. C., Lopez-Amaya, C., Gagne, D., Lovato, L., Moore-Odom, N. H., Kennedy, J.,… Hahn, J. (2010). Anti-myelin antibodies modulate clinical expression of childhood multiple sclerosis. *Journal of Neuroimmunology, 223*, 92-99. doi:10.1016/j.jneuroim.2010.02.019

O'Connor. K. C., McLaughlin, K. A., De Jager, P. L., Chitnis, T., Bettelli, E., Xu, C.,… Dale, R. C. (2007). Self-antigen tetramers discriminate between myelin autoantibodies to native or denatured protein. *Nature Medicine, 13*, 211-217. doi:10.1038/nm1488

Palmer, S. L., Gajjar, A., Reddick, W. E., Glass, J. O., Kun, L. E., Wu, S.,… Mulhern, R. K. (2003). Predicting intellectual outcomes among children treated with 35-40 Gy craniospinal irradiation for medulloblastoma. *Neuropsychology, 17*, 548-555. doi: 10.1037/0894-4105.17.4.548

Palmer, S. N., Meeske, K. A., Katz, E. R., Burwinkle, T. M., & Varni, J. W. (2007). The PedsQL Brain Tumor Module: Initial reliability and validity. *Pediatric Blood and Cancer, 49*, 287–293.

Partington, J. E., & Leiter, R. G. (1949). Partington's Pathway Test. *The Psychological Service Center Bulletin, 1*, 9-20.

Patti, F., Amato M. P., Bastianello, S., Caniatti L., Di Monte E., Ferrazza P.,… Trojano M. (2010). Effects of immunomodulatory treatment with subcutaneous interferon b-1a on cognitive decline in mildly disabled patients with relapsing-remitting multiple sclerosis. *Multiple Sclerosis, 16*, 68–77. doi:10.1177/1352458509350309

Pidcock, F. S., Krishnan, C., Crawford, T. O., Salorio, C. F., Trovato, M., & Kerr, D. A. (2007). Acute transverse myelitis in childhood: Center-based analysis of 47 cases. *Neurology, 68*, 1474-1480. doi:10.1212/01.wnl.000026069.11357.6f

Pidcock, F. S., & Sebire, G. (2011). What is acute transverse myelitis in children? In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 243-254). Cambridge, NY: Cambridge University Press.

Pohl, D. (2009). Epstein-Barr virus and multiple sclerosis. *Journal of the Neurological Sciences, 286*, 62-64.

Pohl, D., Hennemuth, I., von Kries, R., & Henefeld, F. (2007) Pediatric multiple sclerosis and acute disseminated encephalomyelitis in Germany: Results of a nationwide survey. *European Journal of Pediatrics, 166*, 405-412. doi:10.1007/s00431-006-0249-2

Pohl, D., Rostasy, K., Jacobi, Lange, P., Nau, R., Krone, B., & Hanefeld, F. (2010). Intrathecal antibody production against Epstein-Barr and other neurotropic viruses

in pediatric and adult onset multiple sclerosis. *Journal of Neurology, 257*, 212-216. doi:10.1007/s00415-009-5296-y

Pohl, D., & Waubant, E. L. (2011). Treatment of breakthrough disease. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 101-111). Cambridge, NY: Cambridge University Press.

Polliack, M. L., Barak, Y., Achiron, A. (2000). Late-onset multiple sclerosis. *Journal of American Geriatrics Society, 49*, 168, 171.

Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. A., Filippi, M.,… Wolinsky, J. (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, G., 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*, 840-846.

Portaccio, E., Gorretti, B., Lori, S., Zipoli, V., Centorrino, S., Ghezzi, A.,… Amato, M. P. (2009). The brief neuropsychological battery for children: A screening tool for cognitive impairment in childhood and juvenile multiple sclerosis. *Multiple Sclerosis, 15*, 620-626. doi:10.1177/1352458508101950

Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology, 22*, 415-422.

Possa, M. F. (2010). Neuropsychological measures in clinical practice. *Neurological Sciences, 31*, S219-S222. doi:10.1007/s10072-010-0374-6

Ramagopalan, S. V., Morris, A. P., Dyment, D. A., Herrera, B., DeLuca, G. C., Matthew, R.,… Ebers, G. C. (2007). The inheritance of resistance alleles in multiple sclerosis. *PLOS Genetics, 3*, 1607-1613. doi:10.1371/journal.pgen.0030150

Raven, J. C. (1976). *Standard Progressive Matrices*. Oxford, England: Oxford Psychologists Press.

Reitan, R. (1958). The validity of the Trail Making Test as an indication of organic brain damage. *Perceptual and Motor Skills, 8,* 271–276.

Reitan, R. (1969). *Manual for Administration of Neuropsychological Test Batteries for Adults and Children.* Unpublished manuscript, Indianapolis University Medical Center.

Reitan, R. (1971). Trail Making Test results for normal and brain-damaged children. *Perceptual and Motor Skills, 33,* 575–581.

Reitan, R. M., & Davison, L. A. (1974). *Clinical neuropsychology: Current status and applications* (Appendix). Washington, DC: V. H. Winston & Sons.

Renoux, C., Vukusic, S., Mikaeloff, Y., Edan, G., Clanet, M., Dubois, B.,… Confavreux, C. (2007). Natural history of multiple sclerosis with childhood onset. *The New England Journal of Medicine, 356*, 2603-2613.

Renoux, C., & Waubant, E. L. (2011). Pediatric multiple sclerosis course and predictive factors. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the*

central nervous system in childhood (pp. 75-86). Cambridge, NY: Cambridge University Press.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children- Second Edition Manual*. Circle Pines, MN: AGS Publishing.

Richardson, J. T., Robinson, A., & Robinson, I. (1997). Cognition and multiple sclerosis: A historical analysis of medical perceptions. *Journal of the History of the Neurosciences, 6*, 302-319.

Riise, T., Norvtvedt, M. W., & Ascherio, A. (2003) Smoking is a risk factor for multiple sclerosis. *Neurology, 61*, 1122-1124.

Ristori, G., Cannoni, S., Stazi, M. A., Vanacore, N., Cotichini, R., Alfo, M.,… Salvetti, M. (2006). Multiple sclerosis in twins from continental Italy and Sardinia: A nationwide study. *Annals of Neurology, 59*, 27-34.

Rodriguez-Violante, M., Ordonez, G., Bermudez, J. R., Sotelo, J., & Corona, T. (2009). Association of a history of varicella virus infection and multiple sclerosis. *Clinical Neurology and Neurosurgery, 111*, 54-56. doi:10.1016/j.clineuro.2008.07.007

Ronk, F. R., Hooke, G. R., & Page, A. C. (2012). How consistent are clinical significance classifications when calculation methods and outcome measures differ? *Clinical Psychology Science and Practice, 19*, 167-179.

Ronk, F. R., Korkman, J. R., Hooke, G. R., & Page, A. C. (2013). Assessing Clinical Significance of Treatment Outcomes Using the DASS-21. *Psychological Assessment, 25*, 1103-1110. doi:10.1037/a0033100

Ross, K. A., Schwebel, D. C., Rinker, J., Ness, J., & Ackerson, J. (2010). Neurocognitive sequelae in African American and Caucasian children with multiple sclerosis. *Neurology, 75*, 2097-2102.

Ruggieri, M., Iannetti, P., Polizzi, A., Pavone, L., & Grimaldi, L. M. (2004). Multiple sclerosis in children under 10 years of age. *Journal of the Neurological Sciences, 25*, S326-S335. doi:10.1007/s10072-004-0335-z

Ruggieri, M., Plasmati, I., & Simone, I. (2011). Epidemiology of pediatric multiple sclerosis: Incidence, prevalence, and susceptibility risk factors. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 19-35). Cambridge, NY: Cambridge University Press.

Ruff, R. M., & Parker, S. B. (1993). Gender- and age-specific changes in motor speed and eye-hand coordination in adults: normative values for the Finger Tapping and Grooved Pegboard Tests. *Perceptual and Motor Skills, 76*, 1219-1230.

Sadovnick, A. D., & Scheifele, D. W. (2000). School-based hepatitis B vaccination program and adolescent multiple sclerosis. *Lancet, 355*, 549-550.

Salthouse, T. A., Schroeder, D. H., & Ferrer, E. (2004). Estimating retest effects in longitudinal assessments of cognitive functioning in adults between 18 and 60

years of age. *Developmental Psychology, 40*, 813-822. doi:10.1037/0012-1649.40.5.813

Salthouse, T. A. & Tucker-Drob (2008). Implications of short-term retest effects for the interpretation of longitudinal change. *Neuropsychology, 22*, 800-811. doi:10.1037/a0013091

Satz, P. (1993). Brain reserve capacity on symptom onset after brain injury: A formulation and review of evidence for threshold theory. *Neuropsychology, 7*, 273-295.

Sharrack, B., & Hughes, R. A. C. (1996). Clinical scales for multiple sclerosis. *Journal of Neurological Sciences, 135*, 1-9.

Sharrack, B., Hughes, R. A. C., Soudain, S., & Dunn, G. (1999). The psychometric properties of clinical rating scales used in multiple sclerosis. *Brain, 122*, 141-159.

Sherry, A., & Henson, R. K. (2005). Conducting and interpreting canonical correlation analysis in personality research: A user-friendly primer. *Journal of Personality Assessment, 84*, 37-48. doi:10.1207/s15327752jpa8401_09

Shum, D. H., McFarland, K. A., & Bain, J. D. (1990). Construct validity of eight tests of attention: Comparison of normal and closed head injured samples. *Clinical Neuropsychologist, 4*, 151-162.

Simone, I. L., Carrara, D., Tortorella, C., Liguori, M., Lepore, V., Pellegrini, F.,… Livrea, P. (2002). Course and prognosis in early-onset MS: Comparison with adult-onset forms. *Neurology, 59*, 1922-1928.

Slade, P. D., Townes, B. D., Rosenbaum, G., Martins, I. P., Luis, H., Bernardo, M.,… DeRouen, T. A. (2008). The serial use of child neurocognitive tests: Development versus practice effects. *Psychological Assessment, 20*, 361-369. doi:10.1037/a0012950

Smerbeck, A. M., Parrish, J., Serafin, D., Yeh, E. A., Weinstock-Guttman, B., Hoogs, M.,… Benedict, R. H. B. (2011). Visual-cognitive processing deficits in pediatric multiple sclerosis. *Multiple Sclerosis Journal, 17*, 449-456. doi:10.1177/1352458510391689

Smith, A. (1973). *Symbol Digit Modalities Test.* Los Angeles, CA: Western Psychological Services.

Smith, A. (1982). *Symbol Digit Modalities Test (SDMT) Manual-Revised.* Los Angeles, CA: Western Psychological Services.

Snow, W. G., Tierney, M. C., Zorzitto, M. L., Fisher, R. H., & Reid, D. W. (1988). One-year test-retest reliability of selected neuropsychological tests in older adults. *Journal of Clinical and Experimental Neuropsychology, 10*, 60.

Soares, B. P., Chabas, D., & Wintermark, M. (2011). MRI features of multiple sclerosis. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 48-57). Cambridge, NY: Cambridge University Press.

Solari, A., Mancuso, L., Motta, A., Mendozzi, L., & Serrati, C. (2002). Comparison of two brief neuropsychological batteries in people with multiple sclerosis. *Multiple Sclerosis, 8*, 169-176. doi:10.1191/1352458502ms780oa

Speer, D. C. (1992). Clinically significant change: Jacobson and Truax (1991) revisited. *Journal of Consulting and Clinical Psychology, 60*, 402-408.

Spreen, O., & Gaddes, W. H. (1969). Developmental norms for 15 neuropsychological tests age 6 to 15. *Cortex, 5*, 170-191.

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences* (5th ed.). New York, NY: Taylor & Francis Group, LLC.

Strober, L., Englert, J., Munschauer, F., Weinstock-Guttman, B., Rao, S., & Benedict, R. H. B. (2009). Sensitivity of conventional memory tests in multiple sclerosis: Comparing the Rao Brief Repeatable Neuropsychological Battery and the Minimal Assessment of Cognitive Function in MS. *Multiple Sclerosis, 15*, 1077-1084. doi:10.1177/1352458509106615

Sturkenboom, M., Abenhaim, L., Wolfson, C., Roullet, E., Heinzlef, O., & Gout, O. (1999). Vaccinations, demyelination, and multiple sclerosis study (VDMAS): A population-based study in the UK. *Pharamacoepedimiology and Drug Safety, 8*, S170-S171.

Stuss, D. T., Stethem, L. L., & Pelchat, G. (1988). Three tests of attention and rapid information processing: An extension. *The Clinical Neuropsychologist, 2*, 246-250.

Stuss, D. T., Stethem, L. L., & Poirier, C. A. (1987). Comparison of three tests of attention and rapid information processing across six age groups. *The Clinical Neuropsychologist, 1,* 139–152.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness: A practical scale. *The Lancet 2*, 81–4. doi:10.1016/S0140-6736(74)91639-0.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale-Fourth Edition: Technical manual*. Chicago, IL: Riverside.

Till, C., Ghassemi, R., Aubert-Broche, B., Kerbrat, A., Collins, D. L., Narayanan, S.,… Banwell, B. L. (2011). MRI correlates of cognitive impairment in childhood-onset multiple sclerosis. *Neuropsychology, 3*, 319-332. doi: 10.1037/a0022051

Tombaugh, T. N. (2004). Trail Making Test A and B: Normative data stratified by age and education. *Archives of Clinical Neuropsychology, 19*, 203-215. doi:10.1016/S0887-6177(03)00039-8

Tomlinson, D., Hinds, P. S., Ethier, M. C., Ness, K. K., Zupanec, S., & Sung, L. (2013). Psychometric properties of instruments used to measure fatigue in children and adolescents with cancer: A systematic review. *Journal of Pain and Symptom Management, 45*, 83-91. doi:10.1016/j.jpainsymman.2012.02.010

Torisu, H., Kira, R., Ishizaki, Y., Sanefuji, M., Yamaguchi, Y., Yasumoto, S.,… Hara, T. (2010). Clinical study of childhood acute disseminated encephalomyelitis,

multiple sclerosis, and acute transverse myelitis in Fukuoka Prefecture, Japan. *Brain & Development, 32*, 454-462. doi:10.1016/j.braindev.2009.10.006

Trites, R. L. (1977). *Neuropsychological Test Manual*. Ontario, Canada: Royal Ottawa Hospital.

Touze, E., Fourrier, A., Rue-Fenouche, C., Ronde-Oustau, V., Jeantaud, I., Begaud, B., & Alperovitch, A. (2002). Hepatitis B vaccination and first central nervous system demyelinating event: A case control study. *Neuroepidemiology, 21*, 180-186.

Touze, E., Gout, O., Verdier-Tallefer, M. H., Lyon-Caen, O., & Alperovitch, A. (2000). The first episode of central nervous system demyelinization and hepatitis B vaccination. *Revue Neurologique, 156*, 242-246.

Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., & Jolles, J. (2008). Detecting the significance of changes in performance on The Stroop Color-Word Test, Rey's Verbal Learning Test, and the Letter Digit Substitution Test: The regression-based change approach. *Journal of the International Neuropsychological Society, 14*, 71–80. doi:10.10170S1355617708080028.

Van der Mei, I. A., Ponsonby, A., Dwyer, T., Blizzard, L., Simmons, R., Taylor, B. V.,…& Kilpatrick, T. (2003). Past exposure to sun, skin phenotype, and risk of multiple sclerosis: Case-control study. *British Medical Journal, 327*, 316-320. doi:10.1136/bmj.327.7410.316

Varni, J. W., Burwinkle, T. M., Katz, E. R., Meeske, K., & Dickinson, P. (2002). The PedsQL in pediatric cancer: Reliability and validity of the Pediatric Quality of

Life Inventory Generic Core Scales, Multidimensional Fatigue Scale, and Cancer Module. *Cancer, 94*, 2090-2106.

Varni, J. W., Burwinkle, T. M., Limbers, C. A., & Szer, I. S. (2007). The PedsQL as a patient-reported outcome in children and adolescents with fibromyalgia: an analysis of OMERACT domains. *Health and Quality of Life Outcomes, 5*, 9-12. doi:10.1186/1477-7525-5-9

Varni, J. W., Burwinkle, T. M., & Szer, I. S. (2004). The PedsQL Multidimensional Fatigue Scale in pediatric rheumatology: Reliability and validity. *Journal of Rheumatology, 31*,
2494–2500.

Varni, J. W., Limbers, C. A., Bryant, W. P., & Wilson, D. P. (2009). The PedsQL Multidimensional Fatigue Scale in type 1 diabetes: Feasibility, reliability, and validity. *Pediatric Diabetes, 10*, 321-328. doi:10.1111/j.1399-5448.2008.00482.x

Varni, J. W., Seid, M., & Rode, C. A. (1999). The PedsQL: Measurement model for the Pediatric Quality of Life Inventory. *Medical Care, 37*, 126-139.

Venkateswaran, S., Bennett, S., & Ness, J. (2011). Symptomatic therapy in pediatric multiple sclerosis. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 112-133). Cambridge, NY: Cambridge University Press.

Vercellino, M., Plano, F., Votta, B., Mutani, R., Giordana, M.T., & Cavallas, P. (2005). Grey matter pathology in multiple sclerosis. *Journal of Neuropathology and Experimental Neurology, 64*, 1101-1107.

Waldman, A. T., & Balcer, L. J. (2011). Pediatric optic neuritis. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 23-233). Cambridge, NY: Cambridge University Press.

Walker, L. A. S., Mendella, P. D., Stewart, A., Freedman, M. S., & Smith, A. M. (2011). Meaningful change in cognition in multiple sclerosis: Method matters. Canadian *Journal of Neurological Sciences, 38,* 282-288.

Warrington, T. P., & Bostwick, J. M. (2006). Psychiatric adverse effects of corticosteroids. *Mayo Clinic Proceedings, 81*, 1361-1367.

Waubant, E., & Chabas, D. (2009). Pediatric multiple sclerosis. *Current Treatment Options in Neurology, 11*, 203-210.

Waubant, E., Chabas, D., Okuda, D. T., Glenn, O., Mowry, E., Henry, R. G., … Pelletier, D. (2009). Difference in disease burden and activity in pediatric patients on brain magnetic resonance imaging at time of multiple sclerosis onset vs adults. *Archives of Neurology*, 66, 967-971.

Waubant, E., Chabas, D., Strober, J. B., Milazzo, M., Belman, A., & Krupp, L. (2009, May). *25(OH) vitamin D levels in children with pediatric-onset of multiple sclerosis and controls.* Poster session presented at the 14th Annual Meeting of

Americas Committee for Treatment and Research of Multiple Sclerosis, Atlanta, GA.

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children-Revised*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1991). *The Wechsler Intelligence Scale for Children-Third Edition.* San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-Third Edition and Wechsler Memory Scale-Third Edition: Technical manual*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence-Third Edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children-Fourth Edition: Technical and interpretive manual*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale-Fourth Edition: Technical and interpretive manual*. San Antonio, TX: Pearson.

Wechsler, D. (2012). *Wechsler Preschool and Primary Scale of Intelligence: Technical and interpretive manual*. San Antonio, TX: Pearson.

Wingerchuck, D. M., & Weinshenker, B. G. (2000). Multiple sclerosis: Epidemiology, genetics, classification, natural history, and clinical outcome measures. *Neuroimaging Clinics of North America, 10*, 611-623.

Willer, C. J., Dyment, D. A., Risch, N. J., Sadovnick, A. D., & Ebers, G. C. (2003). Twin concordance and sibling recurrence rates in multiple sclerosis. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 12877-12882. doi:10.1073/pnas.1932604100

Williams, J., Rickert, V., Hogan, J., Zolten, A., Satz, P., D'Elia, L., … Light, R. (1995). Children's Color Trails. *Archives of Clinical Neuropsychology, 10,* 211-223.

Wright, B. D., & Stone, M. H. (1979). *Best test design: Research measurement*. Chicago, IL: Mesa Press.

Yeh, E. A., Chitnis, T., Krupp, L., Ness, J., Chabas, D., Kuntz, N., & Waubant, E. (2009). Pediatric multiple sclerosis. *Nature Reviews Neurology, 5*, 621-631. doi:10.1038/nrneurol.2009.158

Yeh, E. A., Parrish, J. B., & Weinstock-Guttmam, B. (2011). Disease progression in pediatric multiple sclerosis: Disparities between physical and neurocognitive outcomes. *Expert Review of Neurotherapeutics, 11*, 433-440.

Yeh, E. A., & Rodriguez, M. (2011). Disease-modifying therapy and response to first-line treatment in pediatric multiple sclerosis. In D. Chabas & E. L. Waubant (Eds.), *Demyelinating disorders of the central nervous system in childhood* (pp. 96-100). Cambridge, NY: Cambridge University Press.

Yeo, T. W., De Jager, P. L., Gregory, S. G., Barcellos, L. F., Walton, A., Goris, A.,… Sawcer, S. (2007). A second major histocompatibility complex susceptibility locus for multiple sclerosis. *Annals of Neurology, 61*, 28-236.

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

Zaaraoui, W., Reuter, F., Rico, A., Faivre, A., Crespy, L., Malikova, I.,… Audoin, B. (2011). Occurrence of neuronal dysfunction during the first 5 years of multiple sclerosis is associated with cognitive deterioration. *Journal of Neurology, 258*, 811-819. doi: 10.1007/s00415-010-5845-4

Zanni, G. R. (2009). Prompt treatment improves outcomes for MS patients. *Pharmacy Times, 75*, 70-73.

Zipp, F., Weil, J. G., & Einhaupl, K. M. (1999). No increase in demyelinating disease after hepatitis B vaccination. *Nature Medicine, 5*, 964-965.

Zivadinov, R., Weinstock-Guttman, B., Hashmi, K., Abdelrahman, N., Stosic, M., Dwyer, M.,… Ramanathan, M. (2009). Smoking is associated with increased lesion volumes and brain atrophy in multiple sclerosis. *Neurology, 73*, 504-510. doi:10.1212/WNL.0b013e3181b2a706

Zorzon, M., Zivadinov, R., Nasuelli, D., Ukmar, M., Bratina, M., Tommasi, M. A.,… Cazzato, G. (2002). Depressive symptoms and MRI changes in multiple sclerosis. *European Journal of Neurology, 9*, 491-496.