The Relationship between Cardiovascular Fitness and Gray Matter Volume in Healthy Adults

by

Andrea M. Weinstein

Bachelor of Arts, University of California, Berkeley, 2005

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of

Arts and Sciences in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

Kenneth P. Dietrich School of Arts and Sciences

This thesis was presented

by

Andrea Weinstein

It was defended on

October 20, 2011

and approved by

Peter Gianaros, Associate Professor, Department of Psychology

Anna Marsland, Associate Professor, Department of Psychology

Thesis Advisor: Kirk Erickson, Assistant Professor, Department of Psychology

The Relationship between Cardiovascular Fitness and Gray Matter Volume in Healthy Adults

Andrea Meriam Weinstein, M.S.

University of Pittsburgh, 2011

Aging is marked by a decline in cognitive function, which is often preceded by losses in gray matter volume. Fortunately, higher cardiovascular fitness (CVF) levels are associated with an attenuation of age-related losses in gray matter volume and a reduced risk for cognitive impairment. Despite these links, we have only a basic understanding of whether fitness-related increases in gray matter volume lead to elevated cognitive function. In this cross-sectional study, we examined whether the association between higher aerobic fitness levels and elevated executive function was mediated by greater gray matter volume in the prefrontal cortex (PFC). One hundred and forty-two older adults (mean age = 66.6 years) completed three classic executive function tasks yielding five measures of ability: incongruent trial reaction time (RT) and percent interference from the Stroop task, number of perseverative errors in the Wisconsin Card Sorting Task (WCST), and forward and backwards digit span length. In addition, participants completed structural magnetic resonance imaging (MRI) scans and CVF assessments. Gray matter volume in the PFC was assessed using an optimized voxel-based morphometry approach. Consistent with our predictions, higher fitness levels were associated with (a) better performance on both the Stroop and WCST tasks, and (b) greater gray matter volume in several regions, including the dorsolateral PFC (DLPFC). Bilateral volume of the inferior frontal gyrus and precentral gyrus mediated the relationship between CVF and Stroop

performance while a non-overlapping region in the left middle frontal gyrus mediated the association between CVF and WCST perseverative errors. Control analyses using a non-prefrontal brain region (right occipital lobe) as a mediator and a task not heavily reliant upon executive function (vocabulary) yielded null results in mediation analyses. This dissociation of brain areas suggests that higher fitness levels are associated with better executive function by means of greater gray matter volume in specific areas of the PFC.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

A combination of better medical care and the aging baby boomer population is causing older adults to become the most rapidly expanding segment of the American population. Between the years 2010 and 2020 there is expected to be a 36% increase in the number of adults over the age of 65 (Administration on Aging 2009). Unfortunately, it is expected that the prevalence of health problems, including cognitive impairment and dementia, will also increase as the population ages. The Alzheimer's Association recently reported that, as of 2010, approximately 5.3 million persons were living with dementia in the United States ("2010 Alzheimer's disease facts and figures," 2010). The costs associated with paying for health and long-term care services for people diagnosed with dementia reached approximately $172 billion in 2010 ("2010 Alzheimer's disease facts and figures," 2010). Such staggering costs associated with dementia beg the need for research to identify effective preventions. Unfortunately, pharmacological treatments for cognitive decline in both healthy and neurologically impaired adults are only minimally effective, prompting the search for other avenues of treatment and prevention. Thus it is an important matter of public health that effective and financially feasible treatments are developed for age-related problems.

1.1 AGE-RELATED CHANGES IN THE BRAIN

Normal aging is associated with structural brain changes that lead to specific declines in cognitive function. The cognitive effects of aging are especially pronounced within the domains of memory and executive functions (Raz, 2000). Executive functions are a set of higher order processes that include cognitive flexibility, planning, goal-oriented behavior, working memory, selecting appropriate and inhibiting inappropriate actions, and high-level abstract reasoning. Many years of research using both patient lesion and neuroimaging studies have confirmed that both the ventral and dorsal areas of the prefrontal cortex (PFC) play a prominent role in the performance of executive function tasks, with dorsal PFC (DLPFC) regions playing a more dominant role over ventral regions (D'Esposito et al., 1995; Miller & Cohen, 2001; Ranganath & Blumenfeld, 2008). Functional neuroimaging and animal studies show that the DLPFC is activated during the maintenance and manipulation of information in working memory (D'Esposito et al., 1995; Petrides, 1995; Ranganath & Blumenfeld, 2008). Flexibly switching between concurrent tasks also relies on DLPFC activation (Badre & Wagner, 2006; D'Esposito et al., 1995; Erickson et al., 2005). Yet, the DLPFC is arguably only one node in a network of regions that support executive function. For example, parts of the parietal cortex, basal ganglia, cerebellum, and hippocampus have also been implicated as parts of neural networks that contribute to executive functioning (Badre & Wagner, 2006; Middleton & Strick, 1994; Woldorff et al., 2004). Unfortunately, all of these regions experience significant tissue loss throughout the lifespan. It is estimated that healthy adults lose approximately 15% of their neocortical tissue between ages 30 and 90, with disproportionately higher losses in areas crucial for executive control (Jernigan et al., 2001; Raz, 2000). In addition, healthy adults over the age

of 55 experience an approximate 1-2% annual decline in hippocampal volume (Raz, Rodrigue, Head, Kennedy, & Acker, 2004). These volumetric decreases of the PFC and parietal brain regions often precede, and lead to, the typical executive function decline seen in cognitive aging (Raz et al., 2005).

Despite this disheartening view of aging, there is promising evidence that cognitive decline may not be as immutable as previously thought. Previous research has found that some people age "successfully" without any significant physical or cognitive problems (Depp & Jeste, 2006; Habib, Nyberg, & Nilsson, 2007; Rowe & Kahn, 1987; Yaffe et al., 2009). Because of this individual variability, research has been able to examine the factors that allow some individuals to age successfully. One factor found from epidemiological studies of successful aging is that greater amounts of physical activity is frequently associated with maintenance of cognitive function across the lifespan (Podewils et al., 2005; Yaffe et al., 2009). For example, in one recent epidemiological study, 2,509 older adults were followed for a period of 8 years. After this follow-up period, participants were divided into those that showed significant cognitive decline, those that showed slower cognitive decline, and those that showed no cognitive decline throughout the eight-year period. The authors of this study examined which factors were more consistently found among those participants that maintained cognitive function over the 8-year period. They reported that 30% of participants maintained cognitive function over the course of 8 years (Yaffe et al., 2009) and weekly amounts of moderate to vigorous exercise increased the odds of maintaining cognitive function in this group by 31%. This finding is encouraging, especially considering that pharmaceuticals intended to combat cognitive decline, such as antioxidant supplements and statins, have proven ineffective as treatments (Masse et al., 2005; Petersen et al., 2005; Sparks et al., 2005; Wolozin, Kellman, Ruosseau, Celesia, & Siegel, 2000).

Similar findings on the relationship between physical activity and reduced risk for cognitive impairment have been reported by several other epidemiological studies and meta-analyses (Andel et al., 2008; Barnes, Yaffe, Satariano, & Tager, 2003; Erickson et al., 2010; Heyn, Abreu, & Ottenbacher, 2004; Podewils et al., 2005; Weuve et al., 2004; Yaffe, Barnes, Nevitt, Lui, & Covinsky, 2001). For example, Podewils and colleagues conducted a prospective study following 3,375 older adult participants over an average of 5.4 years. All participants were assessed for baseline leisure time energy expenditure and engagement in physical activities, while at follow-up participants were assessed for incident dementia. While baseline leisure time energy expenditure was not associated with dementia diagnosis at follow-up, more baseline physical activities was associated with a lower risk for developing dementia after controlling for demographic and lifestyle (e.g. smoking status, alcohol intake, use of hormone replacement therapy) characteristics (Podewils et al., 2005). In fact, individuals who participated in 4 or more physical activities exhibited a 42% decreased risk in developing dementia as compared to individuals who did not participate in any physical activities at baseline. The results from these epidemiological findings support the beneficial impact of physical activity on maintaining cognitive function in late life.

1.2 PHYSICAL ACTIVITY EFFECTS ON THE BRAIN AND COGNITION

Physical activity, such as aerobic exercise, might be both an effective prevention and treatment for cognitive decline. In contrast to most medications, aerobic exercise interventions are consistently associated with increased cognitive performance in older adults (Kramer &

Erickson, 2007; Kramer et al., 1999). Aerobic exercise is generally defined as physical activity that requires increased oxygen consumption to generate energy as compared to rest (Hillman, Erickson, & Kramer, 2008). This is in contrast to anaerobic exercise, which is brief activity that relies only on glucose for energy rather than oxygen (Hillman et al., 2008). The definition of aerobic exercise is slightly different than aerobic fitness, or cardiovascular fitness (CVF), which is the maximum capacity of the cardiovascular system to take in oxygen (Hillman et al., 2008). In general, the term 'aerobic exercise' is used to refer to physical activity *interventions* with an aerobic component, whereas 'aerobic fitness' or 'CVF' is referred to when a single snap-shot of fitness is measured in a cross-sectional or observational design.

To date, aerobic exercise and high CVF have been consistently associated with improved cognitive function across the lifespan, even in adults with AD (Castelli, Hillman, Buck, & Erwin, 2007; Heyn et al., 2004; Kramer et al., 1999). For example, in a seminal study of 124 sedentary adults over age 60, participants were randomly assigned to either a walking group or a control group (Kramer et al., 1999). Both groups attended sessions 3 times per week for 6 months, but the walking group received 45 minutes of monitored and controlled aerobic exercise training whereas the control group received 45 minutes of monitored and instructed anaerobic stretching exercises. The important distinction in this study is that both the experimental and control groups received the same amount of health instruction and social interaction; the groups only differed in that the experimental group received aerobic exercise and the stretching control group received anaerobic exercise. Interestingly, after 6 months, the walking group showed significant improvements in cognition, especially on tasks of executive function, as compared to the stretching control group that showed very minimal improvements in cognition over this period. Overall, this research suggests that executive functioning is disproportionately affected by aging

but that even relatively short bouts (6-months) of aerobic exercise may prove to be an effective method for improving cognitive health in late-life.

Research thus far has established that aerobic exercise improves executive control. Because aerobic exercise appears to exert its strongest effects on measures of executive function, it was hypothesized that brain regions supporting executive function (e.g. DLPFC) would be the most affected by an aerobic exercise intervention. To test this hypothesis, Colcombe and colleagues conducted a two-part study consisting of both cross-sectional and intervention designs to investigate the effect of CVF and exercise on brain function using functional magnetic resonance imaging (fMRI; Colcombe et al., 2004). In study one, a cross-sectional assessment of fitness (CVF) was conducted on older adults and fMRI scans were collected during an attentionally demanding task. Higher-fit participants showed increased activity in the DLPFC and decreased activity in the anterior cingulate cortex during the task compared to their less fit peers. As a follow-up to this, study two examined whether these same differences in brain activity as a function of CVF could be induced by an exercise intervention. Participants were scanned both before and after a 6-month exercise intervention with a similar protocol as that described above (Kramer et al., 1999). Consistent with the results from study one, the exercising participants exhibited a significant increase in activity in the DLPFC and parietal cortex along with a decrease in anterior cingulate cortex activity when compared to the nonaerobic stretching control group. In concert with the neuroimaging results, the aerobically trained individuals were less affected by conflicting contextual information when performing the task than the anaerobic control group. This ability to ignore conflicting contextual information resulted in better scores on the task for aerobically trained and higher fit individuals. The results from this study are important because they establish a clear association between aerobic exercise training, enhanced

executive function, and elevated prefrontal brain activity in a group of older adults that typically show dysfunctional activity in these same brain regions. In short, this study made an important link between exercise training, cognition, and the supporting neural architecture in humans, while at the same time providing an important experimental basis for the results from epidemiological studies.

Physiologically, aerobic exercise improves cognitive performance by inducing the proliferation of neurons and enhancing neural and synaptic plasticity. Experimental studies using rodents show that aerobic exercise increases neuronal survival and neurogenesis, the growth of new neurons, in the dentate gyrus of the hippocampus. For example, in one classic study van Praag and colleagues investigated whether voluntary wheel running would stimulate hippocampal neurogenesis in sedentary mice (van Praag, Christie, Sejnowski, & Gage, 1999). Thirty-four young and old mice were divided into sedentary or voluntary wheel running groups for a period of 2 – 4 months and then sacrificed to examine the degree of cell proliferation in the dentate gyrus of the hippocampus. Notably, voluntary wheel running increased hippocampal neurogenesis in the old group of mice such that they matched the young sedentary group in number of new hippocampal neurons. These effects were seen behaviorally as well: running improved spatial learning acquisition and retention in the Morris water maze for both the young and old mice. These results illustrate several important points. First, they provide a low-level biological justification for the morphological and functional brain results reported in humans. Second, they demonstrate that older adult animals retain the capacity to generate new neurons and that exercise is effective at taking advantage of this plasticity. Third, they show that increases in cell proliferation have direct consequences on learning and memory function in both young and old animals.

Increased cell proliferation requires the production of new vasculature to support the distribution of nutrients and energy to the newly formed cells. Therefore, exercise has also been associated with widespread angiogenesis, or the creation of new vasculature, in the cortex, cerebellum, striatum, and hippocampus (Cotman, Berchtold, & Christie, 2007; Hillman et al., 2008). In the study by van Praag and colleagues described above, voluntary wheel running resulted in cell proliferation, but also resulted in qualitative changes in both surface area and perimeter of blood vessels within the dentate gyrus (van Praag et al. 1999), suggesting that the generation of blood vessels and cells are closely coupled. In another study using 30 minutes of daily forced running, exercising rats showed an increased density of microvessels diffusely throughout the cortex after 3 weeks of training (Ding et al., 2006). This newly created vasculature allows for increased blood flow and nutrient transport to the supported brain areas, resulting in increased brain mass, increased perfusion, and elevated blood volume.

Besides creating new neurons and blood vessels in the brain, rodent research shows that exercise also affects synaptic structure by inducing long-term potentiation (LTP) in the hippocampus. LTP is the synaptic analog for learning and memory formation (Cotman et al., 2007). For example, van Praag et al. (2005) reported that exercising animals show enhanced learning on the hippocampal dependent Morris water maze and enhanced LTP. In fact, enhanced LTP resulting from exercise is directly correlated with enhanced learning on the maze. Other studies have shown similar patterns of electrophysiological indices as a result of exercise (Black, et al., 1990; Farmer et al., 2004; Schmidt-Hieber, Jonas, & Bischofberger, 2004). In sum, animal studies have demonstrated that aerobic exercise not only improves cognitive performance but also enhances the neural circuitry supporting cognitive processing in areas known to be susceptible to age-related deterioration including cortical and hippocampal regions.

The research from animal studies set the stage for examining whether CVF or exercise treatments would be associated with alterations in brain circuitry in humans as assessed by resting state fMRI techniques. For example, increased brain connectivity was demonstrated in a recent aerobic exercise intervention trial in older adults (Voss et al., 2010). Similar to the 6-month interventions described above (Kramer et al., 1999; Colcombe et al., 2004), 65 sedentary participants were randomly assigned to either receive one year of monitored aerobic exercise or one-year of stretching and toning. Three MRI assessments (pre-intervention, after 6-months, and post-intervention) were used to assess how exercise could influence brain connectivity. Consistent with the predictions, older adults in the walking exercise group showed increased functional connectivity between prefrontal, medial-temporal, and lateral occipital areas after one year, as compared to the stretching and toning control group. Furthermore, this increase in functional connectivity was associated with greater improvement in executive function tasks for the exercise group. The results from this study are encouraging since these networks are disrupted in MCI and AD and represent an important focus for treatment outcomes. Furthermore, these results provide a mechanistic explanation for how exercise elicits improved cognition: greater communication and coherence among different brain regions allows for cognitive processing to be more efficient, rapid, and generative.

The research from animal models also provides a basis to examine whether aerobic exercise and CVF could attenuate, or even reverse, brain atrophy in older adult humans. In these studies, structural magnetic resonance imaging techniques and volumetric assessments of the brain are conducted in older adults and assessed as a function of CVF or exercise. For example, using a voxel-based morphometry (VBM) technique to examine cortical volume on a point-by-point basis as a function of age and CVF, Colcombe and colleagues found that, in a group of 55

older adults, higher fit individuals exhibited less age-related gray matter loss in the frontal, parietal, and temporal cortices than lesser fit individuals (Colcombe et al., 2003). Gordon and colleagues replicated this finding in a more recent study (Gordon et al., 2008). That is, they also found that higher fitness levels attenuated age-related losses in medial-temporal, parietal, and frontal areas in a group of 40 older adults. Finally, *in vivo* imaging of cerebral blood volume in the hippocampus replicated the fitness-related angiogenesis effects seen in rodents (Pereira et al., 2007). In this study, eleven healthy adults underwent a 3-month aerobic exercise intervention. Pre- and post-intervention cerebral blood volume maps showed that exercise increased blood flow in the dentate gyrus. Thus, there is growing evidence that aerobic exercise may be a low-cost and highly accessible method for improving cognitive function and reducing cortical atrophy in late adulthood.

1.3 CURRENT STUDY

Past research demonstrates that aerobic exercise influences both functional and structural properties of the PFC. There is now substantial evidence that higher CVF is associated with larger brain volume and better executive function ability (Hillman et al., 2008). However, the extant literature is currently lacking an important link in these associations. That is, we currently do not know whether aerobic exercise improves executive function by affecting the morphology and function of the DLPFC. There has been only one study demonstrating that structural changes in the hippocampus mediate the relationship between CVF and memory performance in older adults (Erickson et al., 2009). Yet research to date has not made an analogous investigation for

the relationship between CVF, PFC volume, and executive functioning. This is a particularly important area of research given the disproportionate atrophy of PFC volume and decline of executive functioning with age (Hillman et al., 2008; Jernigan et al., 2001; Raz, 2000).

The goal of the present study was to determine whether CVF affected executive functioning through an association with gray matter volume in the prefrontal cortex. Specifically, we examined, in a cross-sectional study, whether DLPFC volume mediated the relationship between CVF and executive function in a group of healthy older adults above the age of 60. We used VBM analyses of structural magnetic resonance imaging (MRI) data to determine whether DLPFC volume mediated the relationship between CVF and executive function, as assessed by standardized neuropsychological tests. We hypothesized that PFC volume would mediate this relationship such that higher fitness levels would be associated with larger DLPFC volume, which would then be associated with better executive function. Additionally, we hypothesized that DLPFC volume would specifically mediate the relationship between CVF and executive function, rather than general cognitive performance, given that executive processes are so heavily dependent upon DLPFC circuitry.

2.0 RESEARCH DESIGN AND METHODS

2.1 SUBJECTS

The data for this study was obtained from a sample of 179 participants who participated in the Healthy Active Lifestyle Trial (HALT) at the University of Illinois between 2005 and 2009. The HALT project was a 1-year randomized, controlled trial examining the effect of aerobic activity on the brain and cognition. Data from the HALT project includes socio-demographic information such as age, gender, years of education, race and ethnicity, and occupation; physical information such as health history, height, weight, body mass index (BMI), cardiovascular fitness (assessed by a VO_2 max test), and self-reported physical activity. The HALT project also includes information on cognitive abilities in multiple domains such as general intelligence levels, crystallized intelligence levels, executive functioning abilities, verbal and spatial memory abilities, and reading ability. Finally, the HALT data includes neuroimaging measures, including structural and functional MRI data.

HALT participants were between the ages of 58 and 81 at time of testing (mean age = 66.6 years; standard deviation = 5.6 years). Inclusion criteria for the HALT study were as follows: 60+ years of age during the trial, capability to perform physical exercise, medical consent to perform physical exercise from a personal physician, successful completion of the VO_2 max test, absence of cognitive impairment as assessed by the modified Mini Mental Status

Examination, normal or corrected to normal vision, absence of clinical depression (as measured by the Geriatric Depression Scale; (Sheikh & Yesavage, 1986)), and a sedentary lifestyle at time of baseline assessment. The sedentary lifestyle requirement removes the potential confound that people with better cognitive abilities have a higher propensity to exercise since all participants were at similar levels of lifestyle physical activity. A sedentary lifestyle was defined as participating in no more than 1 twenty minute physical activity per week for the past 6 months, as assessed by the Physical Activity Scale for the Elderly (PASE) (Washburn, Smith, Jette, & Janney, 1993). Cognitive impairment was determined by the modified Mini-Mental Status Examination (Stern, Sano, Paulson, & Mayeux, 1987) and participants were excluded if they did not reach the minimum criteria of 51 points (highest score of 57). This instrument has been verified as having high reliability and validity measures and has been used as a screening tool for dementia in both research and clinical settings (Stern et al., 1987). In addition, all participants met safety criteria for participating in an MRI study. These criteria include no previous history of head trauma, head or neck surgery, diabetes, neuropsychiatric or neurological conditions including brain tumors, or having any ferrous metallic implants that could cause injury due to the magnetic field.

Exclusion criteria for the HALT study included: self-reported regular physical activity (2 or more times per week), physical disability that prohibited mobility, non-consent of a physician, evidence of cardiac abnormalities during the VO_2 max test, clinical depression, presence of implanted devices that would be an MRI safety concern, chronic steroidal treatment, and claustrophobia. Participants with severe asthma or other severe respiratory problems were excluded from the initial HALT trial since these conditions would impair ability to perform the VO_2 max assessment. All participants received a physician's consent to engage in a maximal

graded exercise test (VO_2 max) and signed an informed consent approved by the University of Illinois institutional review board.

The current study included only the cross-sectional baseline information from the HALT participants. It is prudent to initially investigate the relationship between CVF and executive function in a cross-sectional design, rather than proceeding immediately to longitudinal effects. Participant data were excluded from analyses for data quality reasons, including low signal to noise ratio in the MR data, excessive motion during MR data acquisition, structural MR artifacts, and experimental errors in acquiring cognitive assessments. This left 142 participants with sufficient data quality for analysis. CVF was measured by the VO_2 max assessment. Brain volume was examined through VBM analyses of structural MR images and cognitive abilities were assessed through neuropsychological tests.

2.2 INSTRUMENTS AND ASSESSMENTS

2.2.1 Cardiovascular fitness assessment

CVF was assessed by a maximal graded exercise test (VO_2 max) on a motor-driven treadmill. This test assessed maximal rate of oxygen uptake (e.g. maximum volume [V] of oxygen [O_2]). The VO_2 max test is the "gold standard" measurement of CVF (American College of Sports Medicine, 1991). To complete the VO_2 max test, participants walked at a speed slightly faster than their normal walking pace (approximately 3 mph) with increasing grade increments of 2% every 2 minutes, expiring air samples at 30-second intervals. A cardiologist and nurse

continuously monitored measurements of oxygen uptake, heart rate, and blood pressure to ensure patient safety. The test continued until either there was objective evidence that the VO_2 max had been attained (as determined by an exercise physiologist), or the participant volitionally terminated the test due to exhaustion. VO_2 max was defined as the highest recorded VO_2 value when two of three criteria were satisfied: (1) a plateau in VO_2 peak between two or more workloads; 2) a respiratory exchange ratio >1.00; and (3) a heart rate equivalent to their age predicted maximum (i.e. 220 - age). Final VO_2 max scores consist of the maximum gases expired, adjusted for height and weight of the person, measured in units of milliliters per kilogram (ml/kg).

2.2.2 Structural magnetic resonance imaging

All participants underwent structural MRI scanning on a 3 Tesla Siemens Allegra scanner. High resolution T1 weighted brain images were collected using a 3D Magnetization Prepared Rapid Gradient Echo Imaging (MPRAGE) protocol, collecting 144 contiguous slices in an ascending manner (see Erickson et al., 2009 for further scanning details). Scanning paramters included an echo time (TE) of 3.87ms, repetition time (TR) of 1,800ms, field of view (FOV) of 256mm, and an acquisition matrix of 192 x 192 mm.

While the “gold standard” of volumetric studies is the manual tracing of specific areas of interest on high-resolution anatomical images, we used VBM to obtain volumetric brain measurements. There are several problems with manual tracing that limits its usability. First, manual tracing requires extensive time commitments and neuroanatomical knowledge, limiting both the number of subjects and brain regions that can be investigated. It would not be feasible to

conduct manual tracing on a large enough sample size as to warrant sufficient power for a mediation analysis. Second, reproducibility of results is not guaranteed, limiting their generalizability. VBM offers an alternative method to obtain brain volume that is generalizable, semi-automated, reliable, and feasible in a large number of subjects.

VBM allows for a whole-brain volumetric analysis in a semi-automated fashion, making it easily reproducible and accessible to researchers with more varied levels of anatomical knowledge. VBM is feasible and reliable across many populations including healthy older adults (Colcombe et al., 2003; Good et al., 2001), adults with dementia (Guo et al., 2010), diabetic subjects (Musen et al., 2006), and psychiatric populations (Kakeda & Korogi, 2010). VBM analysis computes the probability that each voxel in a structural MR image is cerebrospinal fluid, gray matter, or white matter and yields statistical maps for each voxel type (Ashburner & Friston, 2000). Voxels are then classified into the structural category with the highest probability and can be statistically analyzed between subjects. This approach to analyzing cortical volumes has been well established within the literature. A Pubmed search using the term voxel based morphometry yielded 1,411 studies published in the past 10 years that have used VBM. For volumetric estimates of cortical gray matter in older adults, VBM provides estimates that are similar to manual tracing (Kennedy et al., 2009).

VBM offers several practical strengths that make it a viable analytical method, as compared to other cortical volume estimates such as Freesurfer (http://surfer.nmr.mgh.harvard.edu). For instance, VBM is not platform specific and can be used with several different statistical packages, such as FSL and SPM. Freesurfer, on the other hand, is built on a specific software platform with specific algorithms, limiting the freedom that a researcher has over the analysis. While Freesurfer does yield reliable cortical thickness estimates it is computationally heavy,

requiring approximately 24 hours per subject to analyze data on standard research computer hardware (estimates on https://surfer.nmr.mgh.harvard.edu/fswiki/ReconAllRunTimes). Thus, for the purposes of the current study, VBM was used to estimate cortical gray matter volume.

While there are many advantages to using VBM as the statistical method to determine cortical volume, there are several limitations to this technique as well. VBM analyses provide probability estimates of volume within a voxel, not actual volume values. Indeed, there may be different tissue types included in a single voxel, yet each voxel is classified into a single tissue type. In addition, VBM analyses are conducted on the residual variance that is left after the brain has been warped and transformed into standard space, which can possibly lead to computational problems (Bookstein, 2001). This sequence likely adds some additional error to the volumetric estimates. Additionally, VBM analyzes both cortical thickness and surface area together, which may mask differences that occur between these two dimensions. Despite these limitations, many studies have found consistent and replicable effects using VBM, and VBM is considered an appropriate method to analyze cortical volume in human subjects (Ashburner & Friston, 2001).

2.2.3 Cognitive and lifestyle assessments

2.2.3.1 Stroop task

The Stroop task is a classic test of two key components of executive function: attention and inhibitory control (MacLeod, 1991, 1992). In this task the participant named the color ink in which a word is printed. There are three trial types in the Stroop task: neutral, congruent, and incongruent. Neutral trials are those in which the word is a non-color word (e.g. “Lot”). Congruent trials are those in which the word names the same color as the ink in which it is

printed. Incongruent trials are those in which the word names a different color than the ink. Incongruent trials require the most attentional control and thus provide the most robust measure of executive functioning (Banich et al., 2000). Two outcome measures were obtained from this task: incongruent trial reaction time (RT in ms) and percent interference. Percent interference was calculated as the average incongruent trial RT minus the average congruent trial RT, divided by average congruent trial RT. This interference provided a measure of how much additional executive processing is needed to respond to an incongruent trial above and beyond the processing needed for a congruent trial.

2.2.3.2 Digit span

The digit span subtest of the Wechsler Adult Intelligence Scale- Third Edition (Wechsler, 1997) was administered to each participant. The forward subtest of the digit span measures attentional capacity and short-term memory, whereas the backwards subtest is a measure of working memory capacity (Wechsler, 1997). In this test an experimenter read aloud a series of numbers at an interval of 1 number per second. After completing the series the participant orally repeated the same numbers either verbatim (forward span) or in the reverse order (backward span) of initial presentation. The length of the digit group increased until the participant incorrectly responded to two presentations of the same span length in a row. Span length, or the greatest number of digits correctly manipulated, was used as the outcome measure from this test. The minimum span length score is 0 digits and the maximum span length score is 8 digits. This test has been used extensively throughout clinical and research studies within psychology and is accepted as having high validity and reliability scores among healthy older adults (Ryan & Ward, Jun 1999; Wechsler, 1997).

2.2.3.3 Wisconsin Card Sorting Task (WCST)

Participants completed a computerized version of the WCST, a test which assesses multiple components of executive function including working memory, inhibition, and switching processes (O'Sullivan et al., 2001). The participant's task was to sort cards displayed on a computer screen according to one of three dimensions. The cards contained geometric designs and could be sorted into categories by shape, color, or number of shapes. Participants were asked to match each card that appeared in the lower portion of the computer screen with one of four cards displayed at the top of the screen. The participants were told that the computer would provide feedback about the accuracy of their decision, but that the examiner could not give them any additional instructions about the task. Upon accurate categorization of 10 successive cards the categorization rule changed without participant knowledge. Test discontinuation occurred either when the participant attempted to categorize 128 cards or when 6 categorization rules were correctly completed. The main outcome measure from this task used in the current study is number of perseverative errors. A perseverative error occurred when the participant incorrectly categorized a card by using the same dimensional rule as in the immediately prior trial. This measure is thought to demonstrate cognitive flexibility and switching ability.

2.2.3.4 Vocabulary

This task served as a control cognitive measure to support the hypothesis that the PFC is particularly important for executive functioning rather than cognitive functioning in general. It is a subtest of Ekstrom's Factor-Referenced Cognitive Tests (Ekstrom, French, & Harmon, 1976). In this task, participants viewed 36 different vocabulary words and were to choose the best synonym for that word from four choices. Participants were instructed to skip any words for

which they were completely unsure of since points were deducted for incorrect answers. This test is a measure of verbal comprehension, long-term knowledge, and crystallized intelligence, with reliability calculations ranging between .79 and .94 in adults (Ekstrom et al., 1976). Semantic recognition tasks, while somewhat dependent on PFC integrity, are thought to depend heavily on the hippocampus and perirhinal cortex (Squire, Wixted, & Clark, 2007). In fact, while patients with prefrontal lesions exhibit minor deficits in verbal recognition tasks, overall they perform similarly to healthy controls (Ranganath, 2010; Ranganath & Blumenfeld, 2008). Therefore, PFC volume is not likely to be integral to performance in this task.

2.3 PROCEDURE

Subjects were recruited to the HALT trial through community advertisements and physician referrals. Potential subjects were initially screened over the phone for inclusion and exclusion criteria. Upon passing the initial phone screening, subjects were invited to a group orientation to receive study details and ask questions regarding the program. At the orientation participants were also provided with a battery of psychosocial questionnaires to complete and bring back to the following study visit.

Following these initial visits, subjects completed 4 subsequent baseline sessions. Session 1 took 1 hour to complete a blood draw by a trained phlebotomist and a 24-hour food recall. Session 2 included administration of a neuropsychological battery and a mock MRI session, for a total session length of 1.5 hours. The neuropsychological battery provided cognitive and motor information on each subject and was used to further screen for any signs of dementia. The mock

MRI session ensured that the subject was not claustrophobic and was comfortable with the MRI scanning procedure. Session 3 lasted 1.5 hours and included the cardiovascular assessment of fitness (VO_2 max test). Additionally, subjects completed the physical activity questionnaire (PASE) during this session. Finally, session 4 consisted of the MRI scan, lasting approximately 2 hours. The MRI scan collected structural and functional brain data on each subject.

All baseline sessions were conducted within an approximate timeframe of 2.5 weeks. Combined with the phone screening and group orientation, participants completed all baseline procedures within a 2-month timeline. Baseline sessions were conducted between 2005 and 2009 and data collection ended in 2010.

2.4 STATISTICAL ANALYSES

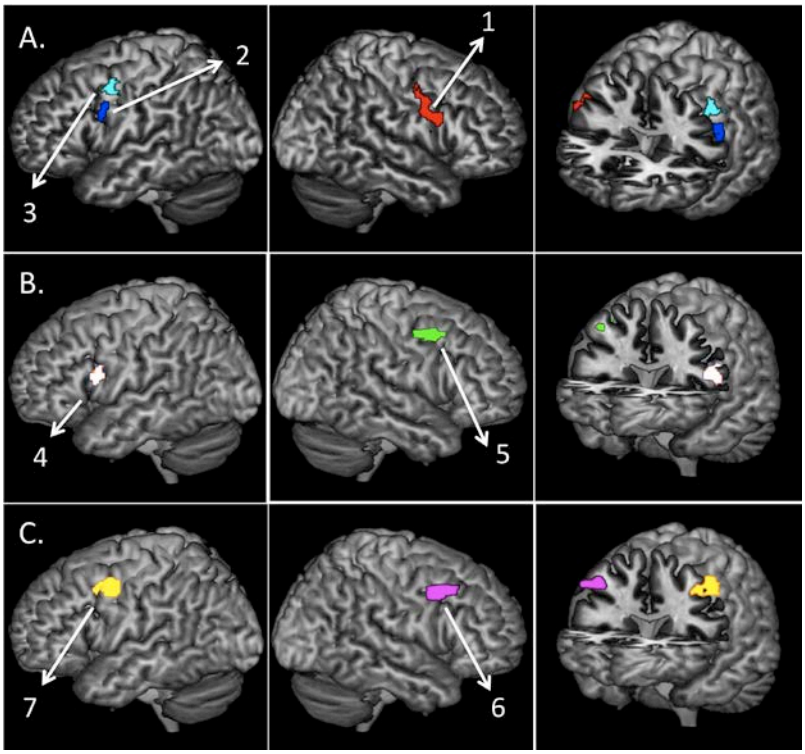
MR data was processed using tools in the FMRIB Software Library (Image Analysis Group, FMRIB, Oxford, UK; http://www.fmrib.ox.ac.uk/fsl/; Smith, Jenkinson et al. 2004). All images were preprocessed to achieve the following results: 1) remove non-brain matter that would obscure statistical analyses, 2) normalize all images into the common Montreal Neurological Institute template, 3) assign volume estimates to each voxel describing the probability that it is either cerebrospinal fluid, gray matter, or white matter, and 4) smooth the tissue maps with a 6.6 mm Gaussian kernel in order to reduce noise. After this segmentation, there was partial volume correction to account for any changes in voxel size due to the normalization step. This allowed for the voxels to represent volumes and not tissue densities.

The DLPFC has been implicated in the monitoring of information in working memory (Petrides, 1995), as well as playing a role in feedback processing during the WCST (Monchi et al., 2001). Attentional control paradigms, such as the Stroop task, are also dependent on DLPFC activation (Banich et al., 2000; D'Esposito et al., 1995; Erickson et al., 2005). Given the prominent role that the DLPFC plays in executive control processes investigated in this study, we focused the initial analyses on the DLPFC so as to reduce potential Type 1 error. The DLPFC was defined bilaterally as the dorsal portions of the superior, middle, and inferior frontal gyri. This definition is consistent with previously validated work by Raz et al. (Raz, Briggs, Marks, & Acker, 1999; Raz et al., 1997).

Using the Threshold-Free Cluster Enhancement (TFCE) technique (Smith & Nichols, 2009), brain regions within the DLPFC were identified that were significantly correlated with CVF when controlling for the variance associated with age, sex, and education. The TFCE technique is completed without using *a priori* thresholds, such as those used in cluster-based thresholding (Smith & Nichols, 2009). Typical cluster and voxel thresholding techniques rely on arbitrarily defined thresholds that can be sensitive to variations in the data near the threshold level. In addition, cluster-based thresholds are sensitive to the amount of spatial smoothing applied before analysis. The TFCE method avoids these limitations by taking the raw (unsmoothed or minimally smoothed) statistical image and producing an output image where each value represents the weighted sum of the local clustered signal (see Smith and Nichols 2009 for further details). TFCE has been validated in several studies and has been shown to be more sensitive than cluster-based or voxel-based thresholding and is the recommended approach for most VBM analyses (Giorgio et al., 2010; Hayasaka, Phan, Liberzon, Worsley, & Nichols, 2004; Salimi-Khorshidi, Smith, & Nichols, 2011; Smith & Nichols, 2009).

Once the DLPFC regions associated with CVF were identified, we next determined regions that were correlated with each cognitive variable. A conjunction between CVF effects on DLPFC volume and each of the cognitive variables' effects on DLPFC volume yielded seven ROIs. Three ROIs were identified based on the conjunction of CVF and Stroop performance (Figure 1a), two from the conjunction of CVF and digit span (Figure 1b), and two from the conjunction of CVF and WCST perseverative errors (Figure 1c). A final region in the occipital lobe was extracted for control analysis. ROI characteristics are described in Table 1. Gray matter volume within these statistically defined regions were extracted and used in a mediation analysis to determine if volume in that region mediated the association between fitness and cognitive performance. The final mediating variable units were partial volume estimates of gray matter.

Figure 1. Regions of interest chosen from conjunction of fitness and cognitive measures



Images are in neurological coordinates and labeled with the corresponding numbers referred to in Tables 1 and 3.

A. Three regions of interest (ROIs) chosen from a conjunction between the main effect of fitness and correlations with Stroop performance.
B. Two ROIs chosen from a conjunction between the main effect of fitness and correlations with Digit Span performance.
C. Two ROIs chosen from a conjunction between the main effect of fitness and correlations with the Wisconsin Card Sorting Task perseverative errors.

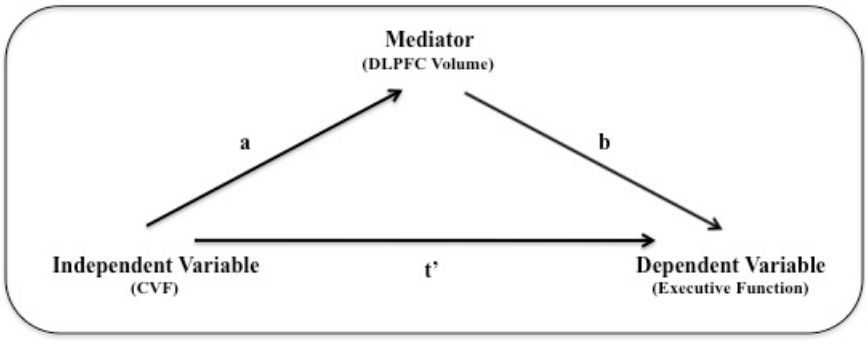
Table 1. Regions of interest chosen for mediation analyses

ROI Description	Number of Voxels	MNI Coordinates (center of gravity)		
		X	Y	Z
1. Right IFG and PCG	122	53	5	33
2. Left IFG and PCG	56	-53	7	32
3. Left MFG	58	-45	2	48
4. Left IFG	159	-51	12	20
5. Right MFG	298	42	1	48
6. Right PCG and MFG	287	44	13	44
7. Left MFG and PCG	328	-41	4	46
8. Occipital Lobe	33	18	-90	4

ROI = region of interest; MNI = Montreal Neurological Institute; IFG = inferior frontal gyrus; PCG = precentral gyrus; MFG = middle frontal gyrus

A mediating variable is a variable that is part of the causal pathway by which an independent variable affects a dependent variable. The main requirement for mediation is that the indirect effect of the independent variable (CVF) through the mediator (DLPFC volume) on the dependent variable (executive function) be significant (Gelfand, Mensinger, & Tenhave, 2009; Zhao, Lynch, & Chen, 2010). This is illustrated as pathways *a* and *b* in Figure 2.

Figure 2. Illustration of the mediation pathway



DLPFC = dorsolateral prefrontal cortex; CVF = cardiovascular fitness

Thus, if pathway *a* → *b* is significant, then DLPFC volume significantly mediates the relationship between CVF and executive functions.

Mediation analyses were conducted using the *indirect* macro designed for SPSS (Preacher & Hayes, 2008). This macro uses bootstrapped sampling to estimate the indirect mediation effect of volume on the relationship between fitness and executive function. In this analysis, 5,000 bootstrapped samples were drawn with replacement from the dataset to estimate a sampling distribution for the indirect mediation pathway. Indirect effects and 95% bias corrected and accelerated confidence intervals are reported. Significant results are indicated by 95% confidence intervals that do not include 0. The dependent variable, executive functioning ability, was computed from each of the five neuropsychological assessment scores: Stroop incongruent trial reaction time and percent interference, forward and backwards digit span lengths, and WCST perseverative errors. Each assessment outcome measure was used as a dependent variable in a separate mediation analysis, rather than a composite executive function index, as these tests measure slightly different aspects of the overall process of executive function and therefore

might be differentially related to DLPFC volume. All regression models controlled for potentially confounding variables that were correlated with the dependent variable, including age, education, and gender. Dummy variables were created for categorical covariates (e.g. gender; reference group = female).

It may be the case that brain volume more generally mediates the relationship between CVF and executive functioning. In order to control for this possibility, a control brain area was assessed for mediation. Given that the occipital lobe is involved in low level visual processing, it is not likely that the occipital lobe is also integral to executive function ability. Thus, a ROI in the occipital lobe was extracted and used for a control mediation analysis. The occipital ROI chosen was constrained to voxels that did not show an association with fitness. This requirement ensured that an area not involved in cardiovascular-related brain changes was chosen for a control region. The mediation analyses were conducted only using the executive function variables in the previous analyses for which DLPFC volume significantly mediated the relationship with CVF. The models were constructed in the same way, except that the occipital lobe ROI was included as the mediator. If occipital lobe volume does not mediate the relationship between CVF and executive function ability, then this provides support for the hypothesis that DLPFC volume specifically, rather than brain volume in general, mediates this relationship.

A final control statistical analysis was conducted to confirm that DLPFC volume specifically mediates the pathway between CVF and executive function ability, rather than CVF and cognitive function in general. A set of mediation analyses identical to those described above was conducted except with vocabulary performance as the dependent variable. While the DLPFC may be involved in verbal comprehension and long-term memory, it is not considered to be the

main brain area necessary for this task. Vocabulary performance is likely to be dependent on areas involved in long-term memory, such as the medial temporal lobe (Ranganath, 2010; Squire et al., 2007). Therefore, it is not likely that the DLPFC mediates the relationship between Vocabulary performance and CVF. If the pathway from CVF to DLPFC volume to Vocabulary performance (i.e. indirect pathway $a \rightarrow b$) is not significant according to the bootstrapped estimates, then there is some support for the conjecture that DLPFC volume specifically mediates the relationship between CVF and executive functioning.

Power analyses indicate that there is sufficient power to detect results from the current study. Using an R^2 of 0.35, as cited by Erickson and colleagues (2009) when regressing hippocampal volume onto CVF in the HALT dataset, only 28 participants would be needed to yield 80% power, well below the current sample size of 142. Gordon and colleagues (2008) found an average R^2 of 0.30 for the relationship between CVF and cognition (WCST perseverative errors and digit span). This effect size indicates that a sample size of only 33 participants would be required to yield 80% power. Finally, in order to determine power for testing the mediation hypothesis, data from a Monte Carlo simulation study comparing different mediation significance testing methods was consulted (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002). Assuming a medium effect size, a sample size between 100 and 200 participants would be needed to achieve 86% power. This sample of 142 participants is well within that range. Given these calculations, there is sufficient power to examine whether DLPFC volume mediates a relationship between CVF and executive functioning.

3.0 RESULTS

3.1 CORRELATIONS BETWEEN CVF AND COGNITION

Characteristics of the 142 participants are described in Table 2. Higher CVF was associated with younger age (r = -.390; p < .001), male gender (F(1,140) = 42.313; p < .001), and higher education (r = .197; p < .05). Consistent with our hypotheses, higher CVF was associated with better performance on tasks of executive function. Specifically, higher CVF was associated with less Stroop percent interference (r = -.195; p < .05) and fewer perseverative errors on the WCST (r = -.172; p < .05). There was no relationship between CVF and Stroop incongruent RT or digit span length (all p's > .05). Accounting for age, gender, and education as covariates in linear regression analyses, higher CVF values predicted less Stroop percent interference (F(4,137) = 1.636; β = -.224; p < .05). CVF did not predict Stroop incongruent RT, digit span length, or WCST perseverative errors after accounting for these covariates.

Table 2. Participant characteristics

Characteristic	Mean (SD)
Age (years)	66.4 (5.5)
Education (years)	15.7 (3.0)
Sex (% female)	64.1%
Cardiovascular fitness level (ml/kg)	21.3 (4.8)
Stroop percent interference	11.5 (12.9)
Stroop incongruent RT (ms)	885.0 (120.5)
Digit span forward (number of digits)	6.5 (1.2)
Digit span backwards (number of digits)	4.9 (1.4)
WCST perseverative errors	19.9 (12.7)
Vocabulary	10.8 (4.4)

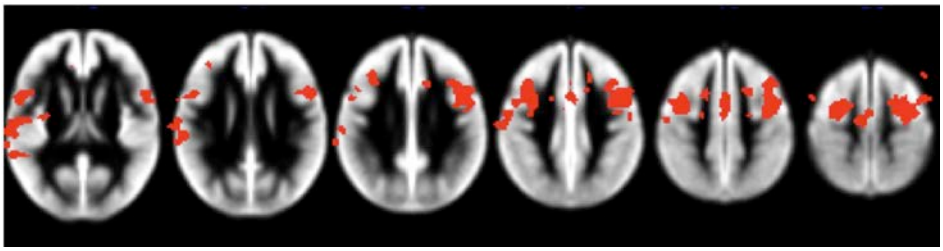
SD = standard deviation; RT = reaction time; WCST = Wisconsin Card Sorting Task

3.2 CVF IS RELATED TO GRAY MATTER VOLUME IN HEALTHY OLDER ADULTS

Consistent with our hypotheses and similar to other studies reporting fitness effects on the brain (e.g. Colcombe et al., 2003, Floel et al., 2010, Gordon et al., 2008), we found that higher CVF levels were associated with greater gray matter volume in several brain regions. Areas in the prefrontal cortex, motor cortex, cingulate gyrus, anterior parietal lobe, and temporal

lobe showed a significant relationship with CVF when controlling for age, sex, and education (Figure 3). These associations remained significant after familywise-error correction of $p < .05$.

Figure 3. Brain regions exhibiting more volume with higher cardiovascular fitness



3.3 THE FITNESS-STROOP PERFORMANCE RELATIONSHIP IS MEDIATED BY DLPFC VOLUME

Mediation analyses were conducted to test the hypothesis that gray matter volume in the DLPFC would mediate the association between CVF and Stroop performance. After controlling for the variance associated with age, sex, and education, gray matter volume bilaterally in the inferior frontal gyrus (IFG) and precentral gyrus (PCG) significantly mediated the relationship between CVF and Stroop incongruent RT (ROIs 1 and 2 in Table 3, shown as the red and dark blue regions in Figure 1a). For the region in the right IFG and PCG, the indirect mediation effect was -1.475 (95% Confidence Interval [CI] = -3.723 : -.162) and for the left IFG and PCG region, the indirect mediation effect was -2.133 (95% CI = -4.604 : -.434). In addition, gray matter volume in the right IFG and PCG significantly mediated the relationship between CVF and Stroop

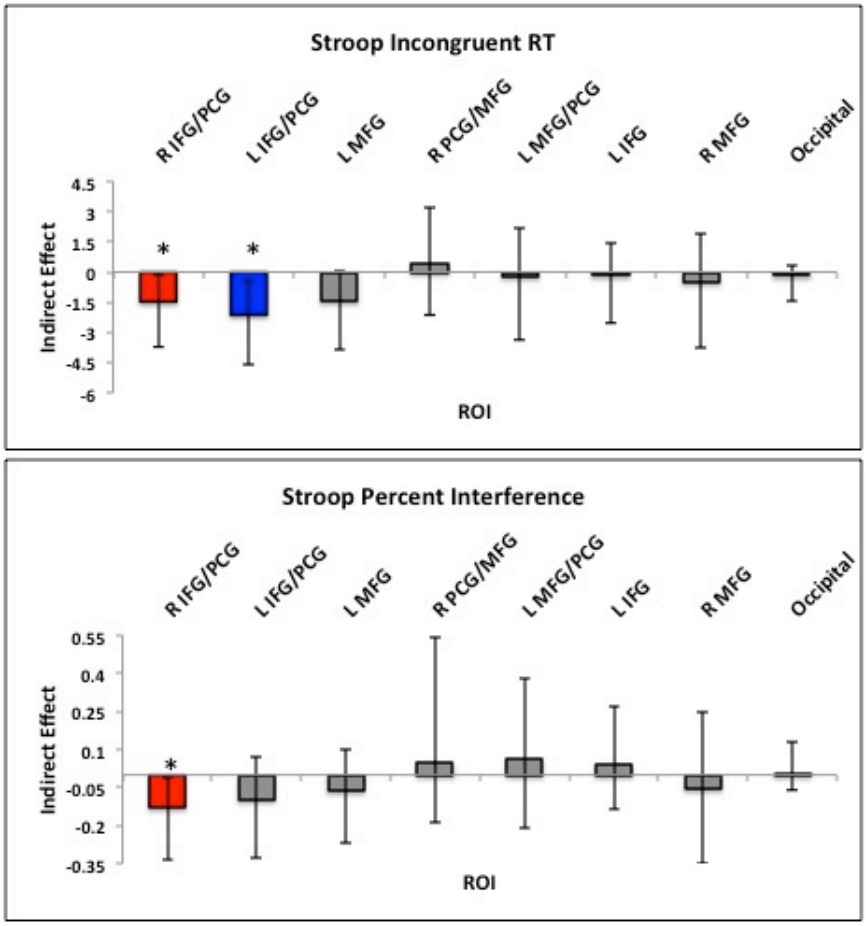
percent interference (ROI 1 in Table 3, shown as the red region in Figure 1a) with an indirect mediation effect of -.128 (95% CI = -.336 : -.011). Mediation effects and 95% CIs are displayed in Figure 4. These effects are negative in direction because the relationships were such that higher CVF was associated with greater gray matter volume, which was in turn associated with shorter incongruent RT and less percent interference, which both indicate better performance on the Stroop task. None of these regions were significant mediators between fitness and WCST or digit span performance.

Table 3. Mediation indirect effects and 95% confidence intervals for each cognitive variable

ROI Description	Stroop Incongruent RT (95% CI)	Stroop Percent Interference (95% CI)	WCST Errors (95% CI)	Digit Span-Forward (95% CI)	Digit Span-Backwards (95% CI)	Vocabulary (95% CI)
1. Right IFG and PCG	**-1.475*** (-3.723 : -.162)	**-.128*** (-.336 : -.011)	.023 (-.115 : .199)	.009 (-.003 : .028)	.011 (-.004 : .032)	.038 (-.003 : .100)
2. Left IFG and PCG	**-2.133*** (-4.604 : -.434)	-.097 (-.327 : .075)	-.041 (-.256 : .131)	.007 (-.008 : .029)	.012 (-.003 : .034)	.037 (-.014 : .117)
3. Left MFG	-1.439 (-3.859 : .070)	-.062 (-.272 : .100)	**-.143*** (-.395 : -.008)	.001 (-.015 : .019)	.010 (-.006 : .033)	.017 (-.034 : .087)
4. Left IFG	-.164 (-2.534 : 1.442)	.041 (-.135 : .267)	-.051 (-.277 : .081)	.008 (-.008 : .034)	.010 (-.007 : .037)	.009 (-.048 : .082)
5. Right MFG	-.492 (-3.767 : 1.917)	-.054 (-.352 : .248)	.101 (-.092 : .423)	.022 (-.006 : .058)	.014 (-.025 : .057)	.013 (-.065 : .109)
6. Right PCG and MFG	.426 (-2.136 : 3.215)	.049 (-.190 : .541)	-.013 (-.403 : .058)	0.000 (-.026 : .038)	.008 (-.031 : .044)	-.005 (-.089 : .094)
7. Left MFG and PCG	-.214 (-3.371 : 2.185)	.062 (-.213 : .381)	-.173 (-.545 : .056)	-.001 (-.025 : .036)	.006 (-.028 : .046)	-.010 (-.104 : .080)
8. Occipital Lobe	-.173 (-1.438 : .290)	.005 (-.060 : .129)	.045 (-.013 : .200)	---	---	---

CI = confidence interval; ROI = region of interest; WCST = Wisconsin Card Sorting Task; IFG = inferior frontal gyrus; PCG = precentral gyrus; MFG = middle frontal gyrus. ***** indicates significant mediation effects determined by CIs that do not include 0.

Figure 4. Mediation indirect effects for Stroop incongruent reaction time and percent interference



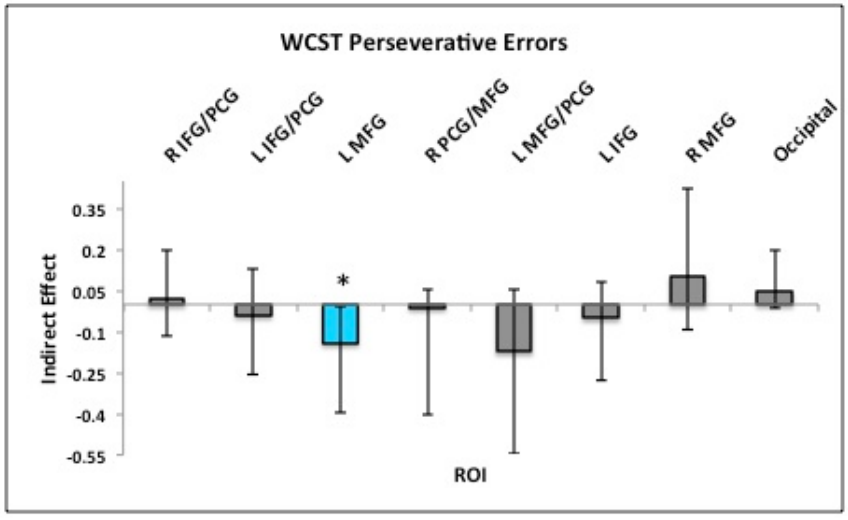
RT = reaction time. Error bars indicate 95% confidence intervals. * indicates significant mediation effects. Colored bars refer to the correspondingly colored regions in Figure 1 that were significant mediators. Regions of interest are named in the same manner as in Tables 1 and 3.

3.4 THE FITNESS-WCST RELATIONSHIP IS MEDIATED BY DLPFC VOLUME

Mediation analyses were conducted to test the hypothesis that gray matter volume in the DLPFC would mediate the association between CVF and WCST perseverative errors. Gray

matter volume in the left middle frontal gyrus (MFG) significantly mediated the relationship between CVF and WCST perseverative errors, when controlling for the variance associated with age, sex, and education (ROI 3 in Table 3, shown as the cyan region in Figure 1a). CVF was associated with greater gray matter volume in this region, which was then related to fewer perseverative errors (indirect effect = -.143; 95% CI = -.395 : -.008; shown in Figure 5). This region only mediated the relationship between CVF and WCST perseverative errors and not any other CVF and cognition relationships.

Figure 5. Mediation indirect effects for the Wisconsin Card Sorting Task perseverative errors

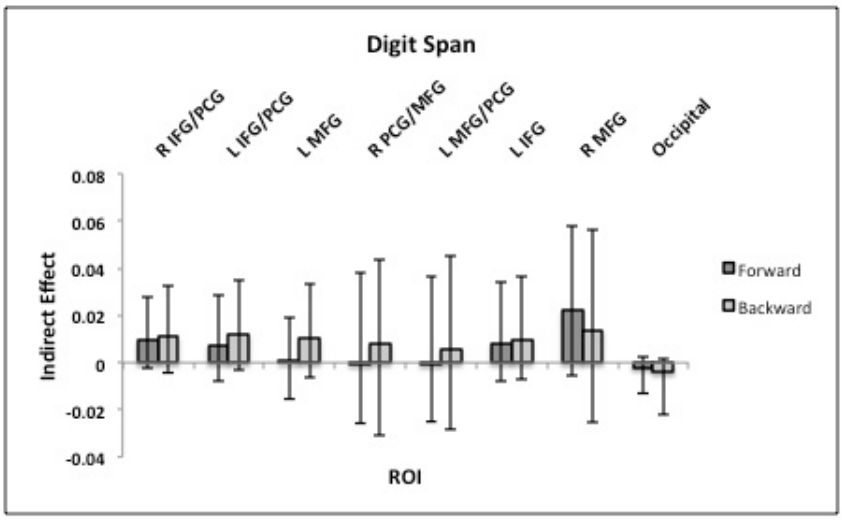


WCST = Wisconsin Card Sorting Task. Error bars indicate 95% confidence intervals. * indicates significant mediation effects. The colored bar refers to the correspondingly colored region in Figure 1 that was a significant mediator. Regions of interest are named in the same manner as in Tables 1 and 3.

3.5 NO RELATIONSHIP BETWEEN FITNESS, DLPFC VOLUME, AND DIGIT SPAN

Mediation analyses revealed no relationship between CVF, DLPFC volume, and either digit span score. Null results are listed in Table 3 and shown in Figure 6.

Figure 6. Mediation indirect effects for the Digit Span task



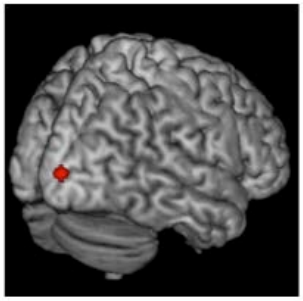
Error bars indicate 95% confidence intervals. Regions of interest are named in the same manner as in Tables 1 and 3.

3.6 CONTROL ANALYSES

Several control analyses were run in order to determine the specificity of the DLPFC mediators. A ROI from the occipital lobe was extracted that did not show an association with CVF (Figure 7). In order to control for the possibility that brain volume generally can significantly mediate a CVF and executive function relationship, we used this occipital ROI as a

mediator in the same analyses conducted above. These control analyses were restricted to only those cognitive variables that already exhibited a significant DLPFC mediation effect: Stroop incongruent RT, Stroop percent interference, and WCST perseverative errors. None of the mediation analyses using this occipital ROI yielded significant results (Table 3).

Figure 7. Occipital lobe region chosen for control analyses



In a final control series, all extracted DLPFC ROIs were used as mediators in a CVF and vocabulary pathway. None of the seven ROIs significantly mediated an association between CVF and vocabulary performance (Table 3). These null control analyses provide evidence for the importance of the DLPFC in the relationship between CVF and executive function and not cognitive function more globally.

4.0 DISCUSSION

Consistent with our predictions and with the extant literature, higher CVF levels were associated with greater gray matter volume throughout much of the cortex, including regions that typically show extensive age-related losses of tissue such as the PFC, cingulate cortex, motor cortex, and temporal lobe (Colcombe et al., 2003; Erickson et al., 2009; Erickson et al., 2010; Erickson et al., 2011; Floel et al., 2010; Gordon et al., 2008). In addition to replicating these prior results, this study demonstrated, for the first time, that greater gray matter volume in the DLPFC mediates the association between aerobic fitness levels and executive function.

Interestingly, a dissociation emerged such that prefrontal brain areas that mediated the CVF association with Stroop performance failed to mediate any CVF association with other cognitive variables (WCST perseverative errors and digit span), while the region that mediated WCST perseverative errors failed to mediate Stroop performance. Bilateral regions in the IFG and PCG mediated a fitness and Stroop relationship, whereas a region in the left MFG mediated a fitness and WCST relationship. All of these mediation pathways were such that higher fitness was associated with greater gray matter volume, which was in turn associated with better cognitive performance (indicated by a negative relationship due to the nature of the variables chosen for analysis). This suggests that higher fitness levels are associated with better executive function by means of greater DLPFC volume. The regions that supported attentional control and

inhibition (Stroop performance) in this sample were located slightly ventral to the region that supported switching ability (WCST perseverative errors).

Most studies (e.g. Colcombe et al., 2003) have argued that greater gray matter volume is beneficial, especially to older adults for whom atrophy is relatively widespread. But how and why greater gray matter volume leads to better cognitive function remains unclear. In fact, the size of particular regions might have little to do with the processing capabilities of that region or the connectedness of the region with other brain areas (Burdette et al., 2010), both of which might be better predictors of cognitive performance than volume. For example, Colcombe et al. (2004) found that task-evoked activity in the prefrontal cortex differed as a function of CVF levels, but these differences occurred independently of any differences in gray matter volume. Similarly, Erickson et al. (2009) reported that greater hippocampal volume mediated the association between fitness and spatial working memory, but that it only accounted for less than 10% of the total variance in spatial memory performance. It may be the case that networks of brain regions are better predictors of cognitive performance than isolated areas (Voss, Erickson, et al., 2010; Voss, Prakash, et al., 2010). Nonetheless, brain atrophy is common in older adults and greater volume is often indicative of less atrophy and better cognitive performance (Raz et al., 2004). Our results support this position since we demonstrate that greater DLPFC volume is associated with better Stroop and WCST performance.

While the molecular pathways by which higher fitness augments DLPFC volume remain unknown, research with animals provides several low-level biological explanations. Physiologically, aerobic exercise enhances learning and memory by inducing the proliferation of cells and by enhancing neural and synaptic plasticity (van Praag et al., 1999). Exercise has also been associated with widespread angiogenesis in the cortex, cerebellum, striatum, and

hippocampus of adult rodents (Cotman et al., 2007; Ding et al., 2006; Hillman et al., 2008). Besides creating new cells and blood vessels in the brain, rodent research shows that exercise also affects synaptic structure by inducing long-term potentiation in the hippocampus (Black et al., 1990; Farmer et al., 2004; Schmidt-Hieber et al., 2004; van Praag, Shubert, Zhao, & Gage, 2005). Exercise may be influencing neural proliferation and vascularization by reducing inflammatory markers or increasing the release of neurotrophins and improving insulin signaling in the brain. Physical activity and fitness increases the production of insulin-like growth factor-1 and brain-derived neurotrophic factor, which facilitate neural and vascular proliferation (Cotman et al., 2007). In addition, insulin can cross the blood-brain barrier and bind to receptors throughout the brain (Carro, Nunez, Busiguina, & Torres-Aleman, 2000; Plum, Schubert, & Bruning, 2005). Insulin inhibits apoptosis and clears β-amyloid from brain tissue, promoting memory formation and overall cognitive health (Plum et al., 2005). This research suggests that aerobic exercise and CVF not only improve cognitive performance but also enhance and expand the neural circuitry supporting cognitive processing in areas known to be susceptible to age-related deterioration. Which of these molecular processes contribute to the enhanced volume found in the current study remains a matter of speculation.

There are several limitations to this study. First, while mediation analyses provide hypotheses regarding pathways of causation, true causal models can only be determined through experimental manipulations. Thus the cross-sectional design of the current study limits any strong causal conclusions. Second, despite including both a control brain region and a control task, it is possible that there is an unmeasured third variable that would improve the explanation of the relationship between CVF and executive function over and above DLPFC volume. For instance, we are not taking into account individual genetic variations that affect the production

and efficacy of neurotrophins. It is also possible that other control variables would be important to include in the analyses, such as mental health history, social network size, and other psychological factors. However, these higher-level models should really only be tested once a formal link between physical activity, brain volume, and cognitive functioning has been firmly established. Third, while conducting an ROI analysis reduces the risk of Type 1 error, this technique leaves open the possibility that a fitness-brain relationship will occur outside of the proposed ROI. Future analyses will be conducted using a voxel-wise whole-brain approach in order to investigate whether any regions outside of the DLPFC, such as the ventrolateral PFC and anterior cingulate cortex, are important in the CVF and executive function relationship. Finally, while VBM provides probabilities of tissue type, these are only estimates. This limitation needs to be taken into account when deriving conclusions from the data. Additional future analyses will use convergent methods to estimate volume, such as Freesurfer, in order to validate significant results. Nonetheless, the current study provides a strong initial investigation of the link between CVF, DLPFC volume, and executive function from which we can subsequently conduct experimental manipulations.

Despite these limitations, this study had several strengths in comparison to previous research. First, the current investigation utilized a sample substantially larger than similar studies in the literature (e.g., Gordon et al., 2008; Peters et al., 2009). This larger sample size allowed us to reliably investigate associations with cognitive performance and to formally test mediation models. In addition, the creation of a study-specific template reduced registration error that often occurs when using atypical populations, such as older adults experiencing age-related atrophy. All subjects were registered to the study specific template using an optimized method that was not available at the time of many previous VBM investigations of fitness-brain relationships (e.g.

Colcombe et al., 2003). In this optimized procedure, all brains were corrected for warping by using the Jacobian determinant, thus reducing registration error even further. Finally, our study used the "gold-standard" metric of CVF, two validated cognitive tasks, and a well-characterized and homogeneous sample to test our hypotheses.

In sum, we have demonstrated, in a large sample of healthy older adults, that higher fitness levels are associated with better cognitive function by the way of increased gray matter volume in the DLPFC. Additionally, it is not the volume of the DLPFC as a whole, but rather specific regions of the DLPFC, that differentially relate to inhibition and switching ability. These results indicate that aerobic fitness influences cognitive function by reducing brain atrophy in targeted areas in healthy older adults.

BIBLIOGRAPHY

2010 Alzheimer's disease facts and figures. (2010). *Alzheimers Dement, 6*(2), 158-194.

Andel, R., Crowe, M., Pedersen, N. L., Fratiglioni, L., Johansson, B., & Gatz, M. (2008). Physical exercise at midlife and risk of dementia three decades later: a population-based study of Swedish twins. *J Gerontol A Biol Sci Med Sci, 63*(1), 62-66.

Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry--the methods. *Neuroimage, 11*(6 Pt 1), 805-821.

Ashburner, J., & Friston, K. J. (2001). Why voxel-based morphometry should be used. *Neuroimage, 14*(6), 1238-1243.

Badre, D., & Wagner, A. D. (2006). Computational and neurobiological mechanisms underlying cognitive flexibility. *Proc Natl Acad Sci U S A, 103*(18), 7186-7191.

Banich, M. T., Milham, M. P., Atchley, R. A., Cohen, N. J., Webb, A., Wszalek, T., et al. (2000). Prefrontal regions play a predominant role in imposing an attentional 'set': evidence from fMRI. *Brain Res Cogn Brain Res, 10*(1-2), 1-9.

Barnes, D. E., Yaffe, K., Satariano, W. A., & Tager, I. B. (2003). A longitudinal study of cardiorespiratory fitness and cognitive function in healthy older adults. *J Am Geriatr Soc, 51*(4), 459-465.

Black, J. E., Isaacs, K. R., Anderson, B. J., Alcantara, A. A., & Greenough, W. T. (1990). Learning causes synaptogenesis, whereas motor activity causes angiogenesis, in cerebellar cortex of adult rats. *Proc Natl Acad Sci U S A, 87*(14), 5568-5572.

Bookstein, F. L. (2001). "Voxel-based morphometry" should not be used with imperfectly registered images. *Neuroimage, 14*(6), 1454-1462.

Burdette, J. H., Laurienti, P. J., Espeland, M. A., Morgan, A., Telesford, Q., Vechlekar, C. D., et al. (2010). Using network science to evaluate exercise-associated brain changes in older adults. *Front Aging Neurosci, 2*, 23.

Carro, E., Nunez, A., Busiguina, S., & Torres-Aleman, I. (2000). Circulating insulin-like growth factor I mediates effects of exercise on the brain. *J Neurosci, 20*(8), 2926-2933.

Castelli, D. M., Hillman, C. H., Buck, S. M., & Erwin, H. E. (2007). Physical fitness and academic achievement in third- and fifth-grade students. *J Sport Exerc Psychol, 29*(2), 239-252.

Colcombe, S. J., Erickson, K. I., Raz, N., Webb, A. G., Cohen, N. J., McAuley, E., et al. (2003). Aerobic fitness reduces brain tissue loss in aging humans. *J Gerontol A Biol Sci Med Sci, 58*(2), 176-180.

Cotman, C. W., Berchtold, N. C., & Christie, L. A. (2007). Exercise builds brain health: key roles of growth factor cascades and inflammation. *Trends Neurosci, 30*(9), 464-472.

D'Esposito, M., Detre, J. A., Alsop, D. C., Shin, R. K., Atlas, S., & Grossman, M. (1995). The neural basis of the central executive system of working memory. *Nature, 378*(6554), 279-281.

Depp, C. A., & Jeste, D. V. (2006). Definitions and predictors of successful aging: a comprehensive review of larger quantitative studies. *Am J Geriatr Psychiatry, 14*(1), 6-20.

Ding, Y. H., Li, J., Zhou, Y., Rafols, J. A., Clark, J. C., & Ding, Y. (2006). Cerebral angiogenesis and expression of angiogenic factors in aging rats after exercise. *Curr Neurovasc Res, 3*(1), 15-23.

Ekstrom, R. B., French, J. W., & Harmon, H. H. (1976). *Manual for the kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Erickson, K. I., Colcombe, S. J., Wadhwa, R., Bherer, L., Peterson, M. S., Scalf, P. E., et al. (2005). Neural correlates of dual-task performance after minimizing task-preparation. *Neuroimage, 28*(4), 967-979.

Erickson, K. I., Prakash, R. S., Voss, M. W., Chaddock, L., Hu, L., Morris, K. S., et al. (2009). Aerobic fitness is associated with hippocampal volume in elderly humans. *Hippocampus, 19*(10), 1030-1039.

Erickson, K. I., Raji, C. A., Lopez, O. L., Becker, J. T., Rosano, C., Newman, A. B., et al. (2010). Physical activity predicts gray matter volume in late adulthood: the Cardiovascular Health Study. *Neurology, 75*(16), 1415-1422.

Erickson, K. I., Voss, M. W., Prakash, R. S., Basak, C., Szabo, A., Chaddock, L., et al. (2011). Exercise training increases size of hippocampus and improves memory. *Proc Natl Acad Sci U S A, 108*(7), 3017-3022.

Farmer, J., Zhao, X., van Praag, H., Wodtke, K., Gage, F. H., & Christie, B. R. (2004). Effects of voluntary exercise on synaptic plasticity and gene expression in the dentate gyrus of adult male Sprague-Dawley rats in vivo. *Neuroscience, 124*(1), 71-79.

Floel, A., Ruscheweyh, R., Kruger, K., Willemer, C., Winter, B., Volker, K., et al. (2010). Physical activity and memory functions: are neurotrophins and cerebral gray matter volume the missing link? *Neuroimage, 49*(3), 2756-2763.

Gelfand, L. A., Mensinger, J. L., & Tenhave, T. (2009). Mediation analysis: a retrospective snapshot of practice and more recent directions. *J Gen Psychol, 136*(2), 153-176.

Giorgio, A., Santelli, L., Tomassini, V., Bosnell, R., Smith, S., De Stefano, N., et al. (2010). Age-related changes in grey and white matter structure throughout adulthood. *Neuroimage, 51*(3), 943-951.

Good, C. D., Johnsrude, I., Ashburner, J., Henson, R. N., Friston, K. J., & Frackowiak, R. S. (2001). Cerebral asymmetry and the effects of sex and handedness on brain structure: a voxel-based morphometric analysis of 465 normal adult human brains. *Neuroimage, 14*(3), 685-700.

Gordon, B. A., Rykhlevskaia, E. I., Brumback, C. R., Lee, Y., Elavsky, S., Konopack, J. F., et al. (2008). Neuroanatomical correlates of aging, cardiopulmonary fitness level, and education. *Psychophysiology, 45*(5), 825-838.

Guo, X., Wang, Z., Li, K., Li, Z., Qi, Z., Jin, Z., et al. (2010). Voxel-based assessment of gray and white matter volumes in Alzheimer's disease. *Neurosci Lett, 468*(2), 146-150.

Habib, R., Nyberg, L., & Nilsson, L. G. (2007). Cognitive and non-cognitive factors contributing to the longitudinal identification of successful older adults in the betula study. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn, 14*(3), 257-273.

Hayasaka, S., Phan, K. L., Liberzon, I., Worsley, K. J., & Nichols, T. E. (2004). Nonstationary cluster-size inference with random field and permutation methods. *Neuroimage, 22*(2), 676-687.

Heyn, P., Abreu, B. C., & Ottenbacher, K. J. (2004). The effects of exercise training on elderly persons with cognitive impairment and dementia: a meta-analysis. *Arch Phys Med Rehabil, 85*(10), 1694-1704.

Hillman, C. H., Erickson, K. I., & Kramer, A. F. (2008). Be smart, exercise your heart: exercise effects on brain and cognition. *Nat Rev Neurosci, 9*(1), 58-65.

Jernigan, T. L., Archibald, S. L., Fennema-Notestine, C., Gamst, A. C., Stout, J. C., Bonner, J., et al. (2001). Effects of age on tissues and regions of the cerebrum and cerebellum. *Neurobiol Aging, 22*(4), 581-594.

Kakeda, S., & Korogi, Y. (2010). The efficacy of a voxel-based morphometry on the analysis of imaging in schizophrenia, temporal lobe epilepsy, and Alzheimer's disease/mild cognitive impairment: a review. *Neuroradiology, 52*(8), 711-721.

Kennedy, K. M., Erickson, K. I., Rodrigue, K. M., Voss, M. W., Colcombe, S. J., Kramer, A. F., et al. (2009). Age-related differences in regional brain volumes: a comparison of optimized voxel-based morphometry to manual volumetry. *Neurobiol Aging, 30*(10), 1657-1676.

Kramer, A. F., & Erickson, K. I. (2007). Effects of physical activity on cognition, well-being, and brain: human interventions. *Alzheimers Dement, 3*(2 Suppl), S45-51.

Kramer, A. F., Hahn, S., Cohen, N. J., Banich, M. T., McAuley, E., Harrison, C. R., et al. (1999). Ageing, fitness and neurocognitive function. *Nature, 400*(6743), 418-419.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychol Methods, 7*(1), 83-104.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychol Bull, 109*(2), 163-203.

MacLeod, C. M. (1992). The Stroop task: The "gold standard" of attentional measures. . *Journal of Experimental Psychology: General, 121*, 12-14.

Masse, I., Bordet, R., Deplanque, D., Al Khedr, A., Richard, F., Libersa, C., et al. (2005). Lipid lowering agents are associated with a slower cognitive decline in Alzheimer's disease. *J Neurol Neurosurg Psychiatry, 76*(12), 1624-1629.

American College of Sports Medicine. (1991). *Guidelines for Exercise Testing and Prescription.* Philadelphia: Lea & Febiger.

Middleton, F. A., & Strick, P. L. (1994). Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science, 266*(5184), 458-461.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu Rev Neurosci, 24*, 167-202.

Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin Card Sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *J Neurosci, 21*(19), 7733-7741.

Musen, G., Lyoo, I. K., Sparks, C. R., Weinger, K., Hwang, J., Ryan, C. M., et al. (2006). Effects of type 1 diabetes on gray matter density as measured by voxel-based morphometry. *Diabetes, 55*(2), 326-333.

O'Sullivan, M., Jones, D. K., Summers, P. E., Morris, R. G., Williams, S. C., & Markus, H. S. (2001). Evidence for cortical "disconnection" as a mechanism of age-related cognitive decline. *Neurology, 57*(4), 632-638.

Pereira, A. C., Huddleston, D. E., Brickman, A. M., Sosunov, A. A., Hen, R., McKhann, G. M., et al. (2007). An in vivo correlate of exercise-induced neurogenesis in the adult dentate gyrus. *Proc Natl Acad Sci U S A, 104*(13), 5638-5643.

Petersen, R. C., Thomas, R. G., Grundman, M., Bennett, D., Doody, R., Ferris, S., et al. (2005). Vitamin E and donepezil for the treatment of mild cognitive impairment. *N Engl J Med, 352*(23), 2379-2388.

Petrides, M. (1995). Impairments on nonspatial self-ordered and externally ordered working memory tasks after lesions of the mid-dorsal part of the lateral frontal cortex in the monkey. *J Neurosci, 15*(1 Pt 1), 359-375.

Plum, L., Schubert, M., & Bruning, J. C. (2005). The role of insulin receptor signaling in the brain. *Trends Endocrinol Metab, 16*(2), 59-65.

Podewils, L. J., Guallar, E., Kuller, L. H., Fried, L. P., Lopez, O. L., Carlson, M., et al. (2005). Physical activity, APOE genotype, and dementia risk: findings from the Cardiovascular Health Cognition Study. *Am J Epidemiol, 161*(7), 639-651.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods, 40*(3), 879-891.

Ranganath, C. (2010). Binding Items and Contexts: the Cognitive Neuroscience of Episodic Memory. *Current Directions in Psychological Science*(19), 131-137.

Ranganath, C., & Blumenfeld, R. S. (2008). Prefrontal Cortex and Memory. In J. H. Byrne (Ed.), *Learning and Memory: A Comprehensive Reference* (Vol. 3, pp. 261-279). Oxford: Academic Press.

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: integration of structural and functional findings. In F. I. M. Craik & T. A. Salthouse (Eds.), *The Handbook of Aging and Cognition* (Vol. 2, pp. 1-90). Mahweh, NJ: Lawrence Erlbaum Associates.

Raz, N., Briggs, S. D., Marks, W., & Acker, J. D. (1999). Age-related deficits in generation and manipulation of mental images: II. The role of dorsolateral prefrontal cortex. *Psychol Aging, 14*(3), 436-444.

Raz, N., Gunning, F. M., Head, D., Dupuis, J. H., McQuain, J., Briggs, S. D., et al. (1997). Selective aging of the human cerebral cortex observed in vivo: differential vulnerability of the prefrontal gray matter. *Cereb Cortex, 7*(3), 268-282.

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., et al. (2005). Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. *Cereb Cortex, 15*(11), 1676-1689.

Raz, N., Rodrigue, K. M., Head, D., Kennedy, K. M., & Acker, J. D. (2004). Differential aging of the medial temporal lobe: a study of a five-year change. *Neurology, 62*(3), 433-438.

Rowe, J. W., & Kahn, R. L. (1987). Human aging: usual and successful. *Science, 237*(4811), 143-149.

Ryan, J. J., & Ward, L. C. (Jun 1999). Validity, reliability, and standard errors of measurement for two seven-subtest short forms of the Wechsler Adult Intelligence Scale—III. *Psychological Assessment, 11*(2), 207-211.

Salimi-Khorshidi, G., Smith, S. M., & Nichols, T. E. (2011). Adjusting the effect of nonstationarity in cluster-based and TFCE inference. *Neuroimage, 54*(3), 2006-2019.

Schmidt-Hieber, C., Jonas, P., & Bischofberger, J. (2004). Enhanced synaptic plasticity in newly generated granule cells of the adult hippocampus. *Nature, 429*(6988), 184-187.

Sheikh, J. I., & Yesavage, J. A. (1986). Geriatric depression scale (GDS): recent evidence and development of a shorter version. In T. L. Brink (Ed.), *Clinical Gerontology: A Guide to Assessment and Intervention* (pp. 165-173). New York: The Haworth Press.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage, 44*(1), 83-98.

Sparks, D. L., Sabbagh, M. N., Connor, D. J., Lopez, J., Launer, L. J., Browne, P., et al. (2005). Atorvastatin for the treatment of mild to moderate Alzheimer disease: preliminary results. *Arch Neurol, 62*(5), 753-757.

Squire, L. R., Wixted, J. T., & Clark, R. E. (2007). Recognition memory and the medial temporal lobe: a new perspective. *Nat Rev Neurosci, 8*(11), 872-883.

Stern, Y., Sano, M., Paulson, J., & Mayeux, R. (1987). Modified mini-mental state examination: Validity and Reliability. *Neurology, 37*, 179.

van Praag, H., Christie, B. R., Sejnowski, T. J., & Gage, F. H. (1999). Running enhances neurogenesis, learning, and long-term potentiation in mice. *Proc Natl Acad Sci U S A, 96*(23), 13427-13431.

van Praag, H., Shubert, T., Zhao, C., & Gage, F. H. (2005). Exercise enhances learning and hippocampal neurogenesis in aged mice. *J Neurosci, 25*(38), 8680-8685.

Voss, M. W., Erickson, K. I., Prakash, R. S., Chaddock, L., Malkowski, E., Alves, H., et al. (2010). Functional connectivity: a source of variance in the association between cardiorespiratory fitness and cognition? *Neuropsychologia, 48*(5), 1394-1406.

Voss, M. W., Prakash, R. S., Erickson, K. I., Basak, C., Chaddock, L., Kim, J. S., et al. (2010). Plasticity of brain networks in a randomized intervention trial of exercise training in older adults. *Front Aging Neurosci, 2*.

Washburn, R. A., Smith, K. W., Jette, A. M., & Janney, C. A. (1993). The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol, 46*(2), 153-162.

Wechsler, D. (1997). *WAIS-III administration and scoring manual*. San Antonio, TX: The Psychological Corp.

Weuve, J., Kang, J. H., Manson, J. E., Breteler, M. M., Ware, J. H., & Grodstein, F. (2004). Physical activity, including walking, and cognitive function in older women. *JAMA, 292*(12), 1454-1461.

Woldorff, M. G., Hazlett, C. J., Fichtenholtz, H. M., Weissman, D. H., Dale, A. M., & Song, A. W. (2004). Functional parcellation of attentional control regions of the brain. *J Cogn Neurosci, 16*(1), 149-165.

Wolozin, B., Kellman, W., Ruosseau, P., Celesia, G. G., & Siegel, G. (2000). Decreased prevalence of Alzheimer disease associated with 3-hydroxy-3-methyglutaryl coenzyme A reductase inhibitors. *Arch Neurol, 57*(10), 1439-1443.

Yaffe, K., Barnes, D., Nevitt, M., Lui, L. Y., & Covinsky, K. (2001). A prospective study of physical activity and cognitive decline in elderly women: women who walk. *Arch Intern Med, 161*(14), 1703-1708.

Yaffe, K., Fiocco, A. J., Lindquist, K., Vittinghoff, E., Simonsick, E. M., Newman, A. B., et al. (2009). Predictors of maintaining cognitive function in older adults: the Health ABC study. *Neurology, 72*(23), 2029-2035.

Zhao, X., Lynch, J. G., Jr., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. *Journal of Consumer Research, 37*, Available at SSRN: http://ssrn.com/abstract=1554563.