2018-06-26

Neuroanatomical Changes Associated with Working Memory Training in Healthy Adults

Savage, Linette

Savage, L. (2018). Neuroanatomical Changes Associated with Working Memory Training in Healthy Adults (Doctoral thesis, University of Calgary, Calgary, Canada). Retrieved from https://prism.ucalgary.ca. doi:10.11575/PRISM/32054

http://hdl.handle.net/1880/106832

UNIVERSITY OF CALGARY

Neuroanatomical Changes Associated with Working Memory Training in Healthy Adults

by

Linette Savage

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN CLINICAL PSYCHOLOGY

CALGARY, ALBERTA

JUNE, 2018

Abstract

The potential for working memory training to enhance cognitive and intellectual abilities is alluring across scientific disciplines and the general public. However, the field has been fraught with inconsistency and controversy. Heterogeneous methodological implementations have led to a divided and contrasting body of literature, which has collectively limited scientific transparency and advancement in the field. However, neuroimaging has the potential to clarify what, if any, benefit working memory training has on the adult human brain. A recent series of studies used functional neuroimaging to investigate neural activations associated with working memory training. This dissertation uses structural imaging to address another theoretical area: the neuroanatomical correlates of working memory training. Forty-eight healthy community dwelling adults, aged 18 - 40 years, completed a series of cognitive tasks and underwent magnetic resonance imaging (MRI) before and after completing a 6-week trial of working memory training (experimental condition) or processing speed training (active control condition). Group by time repeated measures Analyses of Variance (rm-ANOVAs) were conducted on MRI data to identify changes in surface area, thickness, and volume in theoretically relevant gray matter regions of interest, as well as overall gray and white matter volumes, associated with working memory training. Similar analyses were conducted to investigate changes in cognitive task performance in this sample. Null results were present across all neuroanatomical metrics after correction for multiple comparisons, and findings from cognitive tasks were consistent with the subset of literature suggesting that working memory training does not meaningfully benefit cognitive performance. Albeit limited by low statistical power and the confines of available technology, findings of this study, in consort with recently published investigations, strongly support the idea that working memory training is not an effective method for enhancing cognitive performance or inducing neoplastic changes in brain structure. We suggest that future studies continue attempts to resolve heterogeneity and polarization in this field, or alternatively, concentrate resources on identifying and refining mechanisms of change in populations who may benefit from rehabilitative forms of cognitive training.

Keywords: cognitive enhancement, cognitive training, working memory, working memory training, fluid intelligence, n-back training, magnetic resonance imaging (MRI), FreeSurfer

Acknowledgments

I gratefully acknowledge and thank Dr. Vina Goghari for her support throughout my graduate program, including this project. I also thank my dissertation supervisory committee, Dr. Glen Bodner and Dr. Catherine Lebel, for their thoughtful recommendations and continued interest in this project. Thank you to Dr. Lixia Yang and Dr. Frank MacMaster for contributing to this project as part of my examination committee, and Dr. Tavis Campbell for assuming supervisory responsibilities in Dr. Goghari's stead.

Enormous appreciation is extended to those who participated in this study. Thank you for spending your time and mental energy, and I hope you feel you benefited from the process.

Thank you to those who provided the resources for this project. At the Seaman Family Magnetic Resonance Imaging Center, thank you Frances Raymond for juggling the scheduling of numerous projects yet always making me feel that this project was the most important one, and Filomeno Cortese, Dan Pittman, and Jolyn D'Andrea for safely and enthusiastically extracting the best possible neuroimaging data. Thank you FreeSurfer for the MRI analysis software and ongoing support. Thank you Lumosity for providing access to the training games, an intuitive platform for participants, and the comprehensive data generated by your programs. Finally, this work would not have been possible without the generous financial support of the Natural Sciences Engineering Research Council and Alberta-Innovates Health Solutions.

I am eternally grateful to the talented team of individuals who helped with this project. Thanks to Naomi-Rose Dutta for the administrative support. To Aiko Dolatre, Emma Harris, Anna Kazmia, and Averi Howse, your eagerness and commitment to this study, and the oncology study, made this dissertation possible. Thank you to my co-investigator Dr. Cameron Clark. Your endless reserve of intellectual energy and enthusiasm will continue to inspire me.

Reaching this pinnacle is a culmination of many twists and turns. Thank you to my friends who, more than a decade ago, encouraged me as I transitioned from employee to student, especially Leigh who refused my resignation and created a position for me, thereby removing the financial barriers faced by so many other mature students. To Dr. Mike Antle, for creating a classroom and laboratory environment in which I learned things I previously believed I was incapable of learning. To Dr. Linda Carlson who, through your leadership and high standards, taught me how to design, implement, and disseminate research of the highest possible quality. To everyone who shared their knowledge about funding, writing, and grad school in general, thank you! To my lab mates and cohort, you were a fun and supportive group and I am so happy to see your recent successes.

Immeasurable appreciation is reserved for my friends and family who have encouraged and supported me throughout this amazing, frustrating, invigorating, soul-sucking, energizing, impossible, inspiring, terrifying, hilarious, devastating, and awe-inspiring journey. To Alison and Angela, thank you for being voices of support, reason, and silliness. To my mom, you loved and supported me to your last breath. To Brad and Jen, you continue to love and support me, and you created two amazing humans who I look forward to watching grow and thrive. To James, words on paper can never express the love and support you have shown me throughout my education, which roughly corresponds to since the day we met. Thank you for giving me the best of you, and making sure I have the best of everything. Thank you for studying with me, being frustrated for me, laughing with me, pushing me, encouraging me through uncharted territory, and never holding me back. I know you are looking forward to knowing the Linette that is not a student. Sorry to break it to you, I will always be a student. But not a university one, so you do get that.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Nomenclature

Abbreviation	Definition
3T	Three Tesla
3D	Three Dimensional
ANOVA	Analysis of Variance
Aospan	Automated Operation Span
CCFT	Cattell’s Culture Fair Test
CHREB	Conjoint Health Research Ethics Board
DICOM	Digital Imaging and Communication in Medicine
fMRI	Functional Magnetic Resonance Imaging
FOV	Field of View
FSIQ	Full Scale Intelligence Quotient
IPAQ-S7S	International Physical Activity Questionnaire - Last 7 Day, Short Form, Self-Administered
IQ	Intelligence Quotient
mm	Millimeters
MP-RAGE	Magnetization-Prepared Rapid Acquisition Gradient Echo
MRI	Magnetic Resonance Imaging
P-FIT	Parieto-Frontal Integration Theory of Intelligence
POMS-SF	Profile of Mood States - Short Form
PS	Processing Speed
PSQI	Pittsburgh Sleep Quality Index
PST	Processing Speed Training
RAPM	Raven’s Advanced Progressive Matrices
RM-ANOVA	Repeated Measures Analysis of Variance
ROI	Region of Interest
SD	Standard Deviation
SDRT	Spatial Delay Response Task
SPSS	Statistical Package for the Social Sciences
T1	Spin Lattice Relaxation Time
TE	Echo Time
TI	Inversion Time
TR	Repetition Time
WAIS-IV	Wechsler Adult Intelligence Scale - Fourth Edition
WM	Working Memory
WMT	Working Memory Training

Chapter One: Introduction

Over the last decade, researchers and commercial brain training organizations have made bold claims about the benefits of cognitive training. They, either implicitly or explicitly, suggest that cognitive training improves cognitive performance, increases overall intelligence, and slows age or disease related cognitive declines. The implications of successfully enhancing cognitive abilities are broad, spanning multiple scientific disciplines (e.g., cognitive psychology, clinical neuropsychology, gerontology, psychiatry, neuroscience, rehabilitation, education) and populations (e.g., children, adolescents, adults, and older adults who are either healthy or symptom burdened). Therefore, examining claims of the efficacy of cognitive training has value for both the scientific and general community.

Based on their intended purpose, cognitive training investigations fit into three scenarios: remediation (i.e., rehabilitation), inoculation, and enhancement. Cognitive remediation aims to repair, lessen, or compensate for the deleterious cognitive effects of psychiatric conditions, neurodevelopmental disorders, or neurocognitive disorders (see Huntley, Gould, Liu, Smith, & Howard, 2015; McGurk & Mueser, 2017; Tajik-Parvinchi, Wright, & Schachar, 2014). Inoculation approaches are meant to boost cognitive reserve (i.e., existing cognitive capabilities) and brain reserve (i.e., the amount of healthy brain matter available for use by an individual) to prevent future impairment. For example, cognitive training has been used in individuals in prodromal phases of schizophrenia to inoculate them against the negative cognitive effects of the disorder (Rauchensteiner et al., 2011). Individual differences in brain reserve is a proposed explanation for the variance in severity of cognitive symptoms in individuals whose brains are vulnerable to cognitive decline due to disease, injury, or advanced age (Katzman, 1993; Satz, 1993; Stern, 2002). This hypothesis was supported by findings that individuals with larger brain volumes were either less likely to develop the cognitive symptoms that define Alzheimer's disease, or if they did experience symptoms, disease progression was slower (Mortimer, Borenstein, Gosche, & Snowdon, 2005; Mouton, Martin, Calhoun, Dal Forno, & Price, 1998). Such findings support the threshold model of cognitive reserve, which states that increased brain matter volume corresponds with decreased cognitive impairment when brain matter is lost due to disease and/or natural aging processes (Katzman, 1993; Satz, 1993). Consequently, healthy older adults often seek methods to maintain their mental functioning in an effort to stave off future cognitive impairments (Fernandez, 2011). Identifying methods to increase brain reserve may prove useful in preventing or decreasing the negative cognitive consequences in people with vulnerable brains.

The third approach to cognitive training is the enhancement of cognitive functioning and general intelligence in healthy adult populations. In healthy adults, cognitive training refers to any mental activity or structured training program aimed at improving cognitive performance. This broad definition includes training programs focused on enhancing perception, attention, memory (i.e., immediate, recent, and prospective aspects of memory), working memory (i.e., maintenance and manipulation of information in memory), speed of processing, executive functioning, and reasoning. The pragmatic purpose of cognitive enhancement is to increase specific cognitive abilities (e.g., attention, processing speed, working memory, encoding and recall, reasoning) in the hopes that they transfer to broader aspects of day-to-day functioning (e.g., problem-solving, driving ability, remembering details such as names and appointments).

Both cognitive enhancement and cognitive inoculation are the focus of unresolved controversy within the scientific literature. Despite cognitive training's strong conceptual grounding in current models of learning and neuroplasticity, the idea that cognitive training leads to improved cognitive ability and enhanced brain structure remains contentious. Therefore, the purpose of this paper is to 1) review recent findings in the field of cognitive training, with emphasis on acknowledging the controversies, limitations, and theoretical foundations of the field, and 2) build on recent findings to investigate what effect cognitive training has on brain structure. This study will address limitations of previous research and provide neuroimaging data regarding the potential influence of a specific form of cognitive training (i.e., n-back working memory training) on the healthy adult human brain. Findings will contribute to the field of cognitive enhancement, and may support inoculation hypotheses related to cognitive and brain reserve.

1.1 The Brain Training Controversy

Claims regarding the effects of cognitive training have sparked controversy throughout the scientific and general community. In 2014, a collective of 75 scientists, many who authored investigations, reviews, and meta-analyses examining the efficacy of cognitive training programs, signed a consensus statement written to admonish the brain training industry for exaggerated claims (Max Planck Institute for Human Development and Stanford Center on Longevity, 2014). Specifically, they argued, on behalf of the scientific community, that the collective body of research failed to provide compelling evidence that cognitive training is effective, and that commercial brain training marketers were manipulating small, isolated laboratory findings to promote them as significant evidence of the enduring benefits of brain training. For example, one popular commercial brain training provider, Lumosity, was penalized by the Federal Trade Commission for deceptive marketing practices due to unsupported claims that their training games improved work and school performance, protected against dementia, and improved cognitive ability (United States District Court for the Northern District of California, 2016). The consensus group’s primary concerns were that commercial brain training companies were capitalizing on the vulnerabilities of adults wishing to improve or maintain their cognitive faculties (Max Planck Institute for Human Development and Stanford Center on Longevity, 2014).

However, another group of scientists quickly published a rebuttal against the consensus statement. This second group stressed that there was an abundant and growing volume of data, garnered from well-designed investigations, to support the benefits of *some* forms of cognitive training (An open letter to the Stanford Center on Longevity, n.d.). The authors acknowledged that although some findings have been exaggerated for commercial gain, disregarding the findings themselves does a disservice to all stakeholders of cognitive training research, including scientists, funding sources, research participants, and the population in general. In essence, the second group cautioned the first not to reject the scientifically valid positive findings in an attempt to eliminate unwarranted commercially motivated conclusions.

The contrasting positions of these two scientific camps highlight the polarization permeating the field of cognitive training, and its colloquial synonym, brain training. In a commentary about these opposing viewpoints, Merzenich (n.d.) criticized the tendency of scientists to fail to differentiate types of cognitive training programs:

> …using "brain training" as a blanket term is like using "pills" as a blanket term. You would never study one pill and then say "pills" don't work, or do work. Yet researchers (and reporters) do that all the time for brain training programs -- and it makes no more sense than for pills. As with pills, some are highly useful, and others are just sugar. (https://www.cognitivetrainingdata.org/the-controversy-does-brain-training-work/)

Merzenich (n.d.) argued that it is erroneous to coalesce findings from different cognitive training programs and make broad conclusions about the efficacy of cognitive training. Undeniably, heterogeneity is ubiquitous to the field of cognitive training, and others have recognized heterogeneity as a factor leading to different findings in individual studies and in comprehensive reviews and meta-analyses (Melby-Lervag & Hulme, 2013). Therefore, a reasonable approach is to narrow the focus of investigation to a specific form of cognitive training.

1.2 Working Memory Training

Working memory training is a branch of cognitive training that targets cognitive processes within the working memory system. The term working memory represents the cognitive processes involved in the active maintenance and manipulation of discrete amounts of task relevant information in the pursuit of a specific goal (Baddeley, 1992, 2000, 2003, 2017; Baddeley & Logie, 1999; Miyake & Shah, 1999). Although immediate attention and the temporary maintenance of information are necessary aspects of working memory, a defining feature of working memory is its coordinate, executive system referred to as the central executive (Baddeley, 2000; Baddeley & Hitch, 1974). The central executive acts to focus, divide, and switch attention as needed to perceive, maintain, update, and work with the information in immediate storage (Baddeley, 2000; Baddeley & Hitch, 1974). Working memory is a fundamental cognitive ability that underlies higher order thinking (i.e., reasoning through novel information producing dynamic thought, as opposed to reproducing previously learned information) necessary for human abilities such as learning, reading, calculating, creating, conversing, and problem-solving (Baddeley, 2003, 2017; Baddeley & Hitch, 1974; Baddeley & Logie, 1999; Engle, Laughlin, Tuholski, & Conway, 1999; Lewis & Smith, 1993). Given the importance of working memory to a broad range of human cognitive processes, working memory is a valuable target for enhancement.

The basis for hypothesizing that training on a working memory task will enhance working memory and other abilities lies in the properties of the training task itself. By their nature, working memory tasks engage working memory processes. By adapting the task to engage working memory in increasing levels of difficulty, the capacity or efficiency of working memory is thought to increase to meet the demands of the task (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008). Such a theoretical conceptualization is in line with multiple models and theories related to cognitive enhancement and neural plasticity. For example, Thorndike's fundamentals of learning state that repeating an activity will strengthen the connections between the stimulus and the response, particularly when accurate performance is rewarded (Thorndike, 1932). Furthermore, the mismatch model of neural and cognitive plasticity, largely based on Hebbian theory (Hebb, 1949), states that when a discrepancy occurs between available cognitive (e.g., attention, representations, knowledge, strategies) or neural resources (e.g., nuclei structure, dendritic branching, myelination, neurotransmission, vascularization), those resources either increase or decrease to reduce the mismatch (Lövdén, Bäckman, Lindenberger, Schaefer, & Schmiedek, 2010). Training programs that adapt to a trainee's current level of functioning, then progressively advance in difficulty as the participant's ability increases, create a gap between ability and demands, thereby fostering the conditions for plasticity induced cognitive enhancement (Lövdén et al., 2010).

In practical terms, the value of a working memory training program depends on the extent to which training leads to improved cognitive performance beyond the specific task that was trained (i.e., beyond practice effects). Near transfer refers to improvement on a different task within the same cognitive domain that was trained, for example, improvements on one working memory task after training on a different working memory task. Far transfer refers to improvement on a task in a different cognitive domain, such as improvement on a reasoning task after training on a working memory task. Although definitions of near and far transfer vary in the literature (Takeuchi, Taki, & Kawashima, 2010), the value of a transfer effect corresponds to the extent to which the measured outcome is proximal (less valuable) or distal (more valuable) to the trained task.

The hypothesis underlying near and far transfer in most working memory training investigations is that by training and enhancing working memory, higher order cognitive abilities which rely on, or are closely related to, working memory processes ought to be enhanced as a result (Jaeggi et al., 2008; Sternberg, 2008). Fluid intelligence is a frequently investigated far transfer effect because working memory and fluid intelligence share capacity constraints, behavioural mechanisms, and common neural pathways in frontal and parietal brain regions (Burgess, Gray, Conway, & Braver, 2011; Colom, Rebollo, Palacios, Juan-Espinosa, & Kyllonen, 2004; Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Conway, Kane, & Engle, 2003; Gläscher et al., 2010; Gray, Chabris, & Braver, 2003; Halford, Cowan, & Andrews, 2007; Owen, McMillan, Laird, & Bullmore, 2005). Fluid intelligence refers to the ability to solve novel problems through reasoning and without reliance on previously acquired knowledge (Cattell & Cattell, 1959). Fluid intelligence, together with crystalized intelligence (i.e., defined as previously acquired knowledge garnered through the application of fluid intelligence; Sternberg, 2008) are highly correlated with educational attainment, occupational success, economic prosperity, and prosocial behaviour (Gottfredson, 1997). Hence, the potential to enhance fluid intelligence through working memory training holds great allure for cognitive enhancement research, and notable relevance for widespread public benefit.

Numerous reviews and meta-analyses have highlighted the promising effects of working memory training on working memory ability and improved higher order cognitive performance (see Au et al., 2015; Au, Buschkuehl, Duncan, & Jaeggi, 2016; Karbach & Verhaeghen, 2014; Schwaighofer, Fischer, & Bühner, 2015; Weicker, Villringer, & Thöne-Otto, 2016). However, multiple others refute the claim that working memory training provides any meaningful benefit (e.g., Dougherty, Hamovitz, & Tidwell, 2016; Melby-Lervåg & Hulme, 2013, 2016; Melby-Lervåg, Redick, & Hulme, 2016). In general, despite a proliferation of individual studies, reviews, and meta-analyses over the last decade, the issue of whether working memory training is effective remains unresolved and the field remains polarized. Because of this divergence, and the inability for researchers to clarify critical questions related to working memory training (e.g., whether or not working memory training works, and why findings have been so polarized), the field is weakening. Therefore, addressing and resolving the causes of variance is important for moving this research forward. In their comprehensive review and meta-analysis, Melby-Lervåg & Hulme (2013) identified methodological heterogeneity as a primary reason for inconsistent conclusions within the body of working memory training literature. In the section to follow, I present a review of the numerous forms of heterogeneity plaguing the field, and offer suggestions for how these limitations might be overcome.

1.3 The Heterogeneity Problem

Much of the inconsistency among reviews and meta-analyses in the field of working memory training is due to heterogeneity that arises when researchers use different approaches to answer the question: does working memory training work? As demonstrated by the large number of publications that have emerged within the past decade, working memory training remains a dynamic research area with numerous outstanding questions and arguments ripe for investigation, debate, and evolution. However, several aspects of study design and implementation drive differential conclusions. The following specific factors promote divergence in the literature: types of working memory training programs used in individual studies, reviews, or meta-analyses; insufficient sample sizes and diverse sample characteristics; insufficient statistical controls; and different approaches to analyzing and interpreting the same data (Melby-Lervåg & Hulme, 2013; Melby-Lervåg et al., 2016; Moody, 2009; Shipstead, Redick, & Engle, 2012).

1.3.1 Intervention Characteristics

Regarding the training program itself, specific components of an intervention can modulate outcomes and subsequent interpretations. For example, several working memory training programs include immediate memory tasks that only engage the maintenance and simple manipulation processes of working memory, for example, auditory or visual forward and backward span tasks (for examples see www.cogmed.com or cognifit.com). With a focus on simple working memory tasks, these programs fail to recruit complex domain general aspects of working memory such as updating, attentional shifting, inhibition, goal maintenance, and decision-making. Relatedly, span-based working memory training programs permit the development of strategies, which artificially inflate near transfer effects (Morrison & Chein, 2011). Artificially inflated near transfer effects makes far transfer impossible because the working memory system is not actually engaged during training.

In contrast to span-based tasks, complex working memory tasks engage span and more general processes, such as updating and attentional shifting, which define working memory. For example, the n-back task is a cognitively complex task that requires the trainee to shift attention to, maintain, and update auditory and visual stimuli in order to make quick and accurate decisions about the stimuli. As a training task, participants encounter a succession of stimuli and must respond when the current stimulus matches the stimulus observed “n” presentations prior. Variations of the task utilize auditory, visual, or tactile stimuli, or varying combinations thereof. The most frequently used n-back task used in working memory training research is the dual n-back, during which trainees simultaneously receive one visual and one auditory stimulus, and respond when either or both types of stimuli match n-back. The value of n adjusts based on the participant’s success, or difficulty, performing the task. Depending on the trainee’s performance, difficulty can range from n=1 through to n=10. Numerous investigations report beneficial far transfer effects from dual n-back working memory training interventions in healthy adults (e.g., Au et al., 2016; Jaeggi et al., 2010, 2011, 2008; Jaeggi, Buschkuehl, Shah, & Jonides, 2014; Rudebeck, Bor, Ormond, O’Reilly, & Lee, 2012; Schweizer, Hampshire, & Dalgleish, 2011; Stephenson & Halpern, 2013). However, not all dual n-back studies report significant positive far transfer (e.g., Chooi & Thompson, 2012; Colom et al., 2013; Kundu, Sutterer, Emrich, & Postle, 2013; Lilienthal, Tamez, Shelton, Myerson, & Hale, 2013; Redick et al., 2013; Salminen, Strobach, & Schubert, 2012; Seidler et al., 2010; Thompson et al., 2013). The dual n-back task is the subject of investigation in this paper and will be explained in further detail in the methods section.

Despite the distinction between span-based and complex working memory tasks, individual investigations often combine these tasks into a single training program, and reviews and meta-analyses commonly lump together the effect sizes of both forms of training (Weicker et al., 2016). Weicker and colleagues (2016) observed that although simple span interventions contribute to variability in working memory outcome measures ($\chi^2 = 110.08$, $p < .05$, $I^2 = 25\%$), even more variability can be accounted for ($\chi^2 = 62.97$, $p < .05$, $I^2 = 44\%$) by working memory training programs that target domain general processes. Problematically, reviews and meta-analyses evaluate both types of training as if they were similar, despite the fact that span based training programs do not adequately engage working memory (Cowan et al., 2005; Mccabe et al., 2010; Morrison & Chein, 2011; Takeuchi et al., 2010).

Moreover, some working memory training programs combine working memory tasks with tasks falling within other cognitive domains (e.g., reasoning, problem-solving), which results in an overestimate of far transfer (e.g., Hardy et al., 2015). Alternatively, overly inclusive programs and study designs that lack methodological rigour can lead to sweeping claims that training is ineffective (e.g., Owen et al., 2010). In sum, including different forms of training, which produce different effects, under the umbrella of working memory training, is bound to result in inconsistent conclusions.

1.3.2 Sample Characteristics

Another major issue leading to inconsistent conclusions involves the analysis of small and heterogeneous samples. With few exceptions, sample sizes in working memory training studies are small (i.e., less than 20 per group) calling to question the influence of error on significant findings (Melby-Lervåg & Hulme, 2013; Morrison & Chein, 2011; Shipstead et al., 2012). Underpowered studies may either fail to detect effects, or lead to increased statistical variability resulting in overestimated effects. Conversely, the few studies that have large sample sizes (e.g., Owen and colleagues, N = 11,430 and Hardy and colleagues, N = 4,715) do so by sacrificing methodological rigour. Specifically, these studies are conducted fully online with little control over who is completing the training and outcome measures (e.g., minimal exclusion criteria, no monitoring of whether the study participant is the individual completing training and outcome tasks), for how long (e.g., no monitoring of compliance resulting in a wide range of training dosage), and under what conditions (e.g., non-standardized administration of cognitive outcome tasks). Conversely, well controlled studies that follow strict methodological protocols (e.g., standardized, in-person cognitive measurement; compliance monitoring throughout training) are limited by statistical power due to resource limitations (e.g., time and cost to train and test each participant).

Adding to complications surrounding sample size is the tendency for reviews and meta-analyses to combine a variety of sample characteristics in their analyses. For example, prominent meta-analyses questioning the validity of working memory training are composed of studies in which participants range in age from four years through to 70 years and above (Melby-Lervåg & Hulme, 2013; Weicker et al., 2016). Relative to younger adults (age 19-28), middle-aged adults are less adept at dual n-back tasks, presumably related to memory load requirements (Jaeggi, Schmid, Buschkuehl, & Perrig, 2008). Hence, it is possible that older participants benefit more from training by increasing such memory load requirements. Conversely, given that plasticity decreases with age, the potential for change in younger participants may be greater than for participants approaching middle-age. Such age-related factors make it important to investigate discrete age ranges, or at least include age as a potential factor related to changes after working memory training.

Furthermore, most meta-analytic studies aggregate results from healthy participants, and those with various conditions that impact cognitive performance, including learning disabilities, neurodevelopmental disorders (e.g., Attention Deficit - Hyperactivity Disorder, Intellectual Disability), and neurological conditions such as stroke and acquired brain injuries (Melby-Lervåg & Hulme, 2013; Weicker et al., 2016). Relative to meta-analyses with a more restrictive sample of healthy adults (e.g., Au et al., 2015), studies with heterogeneous samples have lower overall effect sizes (e.g., Melby-Lervåg & Hulme, 2013, 2016; Schwaighofer et al., 2015). In line with

Fletcher's (2007) commentary, it does not make sense to compare such varied samples because the emergent results are less able to predict specific outcomes for discrete samples.

1.3.3 Statistical Controls

Various types of control groups are found within the collective of working memory training studies, leading to variable effects. Numerous studies rely on passive control groups to control for practice effects; however, passive control groups give rise to larger effects relative to studies with active control conditions (Melby-Lervåg & Hulme, 2016). Problematically, uncontrolled studies are more likely to result in the dissemination of positive effects; whereas, actively controlled studies result in more conservative estimates.

Moreover, too much similarity between the training and active control tasks can modulate effect sizes. This view is supported by studies that failed to detect significant effects of working memory training, likely because tasks in the training and control conditions both recruited adaptive visual spatial attention and processing (see Kundu et al., 2013; Redick et al., 2013; Thompson et al., 2013). In these studies, training tasks may have been too similar to allow for differential training effects to occur. This idea is further exemplified by Stephenson and Halpern's (2013) study in which dual n-back training was compared to four other conditions: auditory only n-back training, visual only n-back training, visual short-term memory training, and a no-training control group. Similar transfer effects were identified in the three visual task training groups, relative to the auditory or no-training control groups. Collectively, these findings suggest that visual memory tasks, when included in control conditions, likely mask true effects of working memory training.

In sum, the presence or absence of an active control group influences the differential yield of a training program. Furthermore, the characteristics of the control procedure itself can influence the strength of measured effects. Consequently, when control group quality and composition are inconsistent across studies, as is the case in the field of working memory training, variable effects and inconsistent conclusions are bound to emerge.

1.3.4. Interpretation of Outcomes

Further to variable methodologies involving characteristics of training programs, control groups, and study samples, heterogeneity in the measurement of data and interpretation of results leads to different conclusions being drawn from the same data. A significant analysis and discussion of this problem was presented by Redick (2015) who questioned the interpretations of a group of studies claiming positive effects of working memory training. Specifically, in one study, the control group scored higher on baseline outcomes relative to the training group, then their scores decreased after training, creating an effect driven by baseline differences and control group changes (Schweizer et al., 2011). Similarly, presumably by chance, some control groups score lower after training relative to before training (for examples see Roughan & Hadwin, 2011; Vartanian et al., 2013; Zinke et al., 2014). In these studies, the data were interpreted to support the effects of working memory training, despite the likelihood that decreased scores in the control groups, rather than improvement in the experimental groups, drove interaction effects (Redick, 2015).

Another problem in the field involves inappropriate measurement of the transfer effects of interest. Specifically, some authors report far transfer effects despite notable congruence between the working memory training task and the far transfer measure (e.g., Hardy et al., 2015). Others rely on single measures of the far transfer construct of interest, a practice that increases variability resulting in either overestimation or underestimation of effects. Specifically, the seminal study published by Jaeggi and colleagues (2008) which revealed dose-dependent benefits of n-back training to near and far transfer, was criticised for utilizing different fluid intelligence measures for different dosage conditions. More importantly, Jaeggi and colleagues' (2008) decision to half the time allotted for participants to solve fluid intelligence puzzles resulted in a task that measured working memory (near transfer) rather than fluid intelligence (far transfer; Moody, 2009). The use of single outcome measures is particularly concerning when the single measure is flawed, a concern expressed by Moody (2009) in a commentary about measuring fluid intelligence in working memory training research. However, even when two measures of a far transfer construct are included, authors have been known to interpret findings in a direction that supports their hypotheses, even when further examination of the same data indicates that effects are not consistent and therefore not robust enough to support positive conclusions about transfer effects (Lawlor-Savage & Goghari, 2016; Savage, 2013).

Similarly, variability in the interpretation of the same meta-analytic data causes confusion within the field. For example, Au and colleagues' (2015) meta-analysis of dual n-back studies supported far transfer effects, but was soon followed by Melby-Lervåg and Hulme's (2016) criticism refuting Au and colleagues' (2015) reported interpretation of findings. In response to this criticism, Au and colleagues (2016) reanalyzed their data and defended their original conclusions. However, Dougherty and colleagues (2016) re-analyzed the same data with a Bayesian approach, and results failed to support far transfer effects of n-back working memory training.

In sum, different approaches to data analysis lead to different findings and continued confusion about the efficacy of working memory training. Similarly, the promotion of data in specific ways perpetuates scepticism and polarization within the field. Therefore, the field is in need of unbiased, objective forms of measurement, analysis, and dissemination.

1.3.5. Resolving the Heterogeneity Problem

As noted by Morrison and Chein (2011), methodological inconsistencies are a major obstacle to the field because investigators are unable to conclude whether conflicting findings are due to differential effects of specific training programs, dissimilar experimental practices, or individual differences in age or health status present within the study sample. Further complicating the field is the tendency for reviews and meta-analyses to coalesce a diverse mix of participants (e.g., healthy, cognitively impaired, young, old) and training program types into a single meta-analysis, which blunts the effects of individual studies and clouds overall meta-analytic conclusions. This problem becomes compounded because reviews and meta-analyses often rely on vote-counting techniques, in which each significant effect is considered a positive and each non-significant effect a negative, even when studies are underpowered (Karbach & Verhaeghen, 2014). Variable methodology results in variable effect sizes (Morrison & Chein, 2011), thus equalizing the "votes" counted in meta-analyses and reviews resulting in a polarized body of literature.

In order to address the methodological inadequacies present in working memory training studies, Buschkuehl and Jaeggi (2010) and Melby-Lervåg and Hulme (2013) described the features of an ideal working memory training study. Based on their recommendations, the ideal working memory training study ought to include adequate sample sizes (e.g., a minimum of 20; Simmons, Nelson, & Simonsohn, 2011), randomized assignment to groups, and multiple measures of the far transfer outcome of interest. A further important recommendation is the presence of an active control group, which, should an investigator be required to choose between the two due to resource limitations, is of greater priority than the inclusion of a passive control condition (Melby-Lervåg & Hulme, 2013). When meta-analyses can be conducted using an adequate number of well-designed and well-powered studies, based on a homogenous training program and sample, clarity may begin to emerge in the field. However, Buschkuehl and Jaeggi (2010) note that,

> …although it is easy to conceptualize the ideal study, one has to keep in mind that the logistics for running an intervention study are usually very challenging, expensive, and time consuming. Therefore, progress may not take place as quickly in this field as it does in others. (pp. 270)

1.4 Theoretical Rationale for Working Memory Training and Transfer

In addition to resolving the heterogeneity problem, the field of working memory training will benefit from the presentation of a clear theoretical rationale. Most studies focus on whether or not working memory training works, yet many do so without outlining the ideas and assumptions (e.g., potential mechanisms) underlying the research (Redick et al., 2013). Studies that do address the theoretical rationale of working memory training state that, as a function of working memory and associated neural correlates, distinct but related cognitive abilities and neural processes can be improved (Jaeggi, et al., 2008; Sternberg, 2008). The general idea is that when working memory abilities are enhanced, higher order cognitive abilities such as fluid intelligence, or even intelligence more generally, will be enhanced via the mechanisms shared among the processes of working memory that also underlie higher order abilities.

Fluid intelligence is a frequently investigated target of far transfer due to its notable concordance with working memory and its importance to more general aspects of functioning. As noted previously, working memory and fluid intelligence share behavioural mechanisms (e.g., shared variability ranging from .3 to .9; Burgess et al., 2011; Colom et al., 2004) and capacity constraints (e.g., approximately four items for working memory, or four interrelated variables for fluid intelligence; Halford et al., 2007). Three specific processes of working memory are most often cited as sharing variability with fluid intelligence: updating, attentional shifting, and inhibition; however, updating appears to stand out from the other two processes as a particularly important shared mechanisms between working memory and fluid intelligence (Friedman et al, 2006). Furthermore, working memory and fluid intelligence share neural mechanisms within frontal and parietal brain regions (Burgess et al., 2011; Clark, Lawlor-Savage, & Goghari, 2017a; Conway et al., 2002; Gläscher et al., 2010; Gray et al., 2003; Halford et al., 2007; Owen et al., 2005).

Regarding neural mechanisms of working memory, functional neuroimaging studies strongly indicate activation of frontal and parietal cortical regions during performance of working memory tasks (Jolles, Van Buchem, Crone, & Rombouts, 2013; Klingberg, 2010; Kundu et al., 2013; Olesen, Westerberg, & Klingberg, 2004; Owen et al., 2005). This idea was exemplified by a comprehensive meta-analysis of activation coordinates reported across 24 studies utilizing an n-back task. Specifically, Owen and colleagues (2005) identified five consistently activated bilateral cortical regions: dorsolateral prefrontal cortex, medial posterior parietal cortex, premotor cortex, dorsal cingulate / medial premotor cortex including supplementary motor area, rostral prefrontal cortex, and mid-ventrolateral prefrontal cortex. A more specific consideration of particular task variants indicated greater activation in the left ventrolateral prefrontal cortex during verbal n-back tasks relative to non-verbal stimuli (e.g., shapes), and greater activation in the right dorsolateral prefrontal, lateral premotor, and posterior parietal cortex for non-verbal location monitoring n-back tasks (Owen et al., 2005). The authors noted that verbal task activations were consistent with neural areas related to inner speech, while spatial task activations were consistent with neural areas associated with spatial processing (Owen et al., 2005). Similar evidence of the functional underpinnings of working memory emerged from a study that measured activations before and after visuospatial and span-based working memory training. Olesen and colleagues (2004) identified increased activity in prefrontal and parietal areas (specifically, middle frontal gyrus and the superior and inferior parietal cortices) after five weeks of adaptive training.

Overlapping areas are implicated in non-verbal (i.e., fluid) intelligence tasks, and intelligence more generally. In lesion studies, right dorsolateral prefrontal cortex lesions (Barbey, Koenigs, & Grafman, 2013) and posterior parietal lesions (Waechter, Goel, Raymont, Kruger, & Grafman, 2013) were associated with impairments in fluid reasoning ability. Neuroimaging studies have consistently found increased recruitment of prefrontal cortical areas during the performance of fluid intelligence tasks (Christoff et al., 2001; Duncan & Owen, 2000; Kroger, 2002; Prabhakaran, Smith, Desmond, Glover, & Gabrieli, 1997). Furthermore, findings of a comprehensive review of biological markers of intelligence led to the Parieto-Frontal Integration Theory (P-FIT) of intelligence, which suggests that variations in specific frontal and parietal regions, namely the dorsolateral prefrontal cortex, the inferior and superior parietal lobule, and the anterior cingulate, explain variations in human reasoning and intelligence task performance (Jung & Haier, 2007). The P-FIT theory of intelligence has also been supported by neuroanatomical investigation. Specifically, Colom and colleagues (2009) identified gray matter volumes correlated with intelligence in expected frontal and parietal areas, specifically, the dorsolateral prefrontal cortex, Broca's and Wernicke's areas, and in somatosensory and visual association cortices.

In sum, working memory and fluid intelligence share key cognitive and neural mechanisms, and positioned within the framework of the P-FIT theory of intelligence, working memory emerges as a reasonable target of cognitive training when the goal is to effect change in fluid reasoning. However, despite a reasonable conceptual rationale, the answer to the question of whether working memory training has any meaningful effect on cognitive processes or brain structure and function remains elusive.

1.5 Investigating Structural Brain Plasticity

Despite the large body of working memory training studies that have attempted to measure performance changes on psychometric tasks after working memory training, findings remain inconsistent and conflicting. However, including neuroimaging as a measure bypasses some of the limitations of behavioural studies. Most notably, neuroimaging can provide an objective measure of training related effects, which may occur irrespective of changes being detected on cognitive tasks (Colom et al., 2016; Engvig et al., 2010; Román et al., 2016). Moreover, investigating and identifying training induced structural change has implications for challenging existing theories and contributing to cognitive enhancement, inoculation, and rehabilitation approaches. Similarly, by identifying anatomical changes in neural areas associated with the training and far transfer task (e.g., working memory and fluid intelligence), investigators can relate findings directly to the assumptions underlying cognitive training, namely, that training has some effect on the brain, and specifically, that training has an effect on brain metrics important for cognitive proficiency and longevity.

Studying neural macrostructure contributes to scientific knowledge about experience dependent neural plasticity in healthy adults, which remains a relatively new field. Only within the last twenty years has the traditionally held view, that neuroanatomical changes are restricted to critical developmental periods, shifted to an acceptance that the human brain can adjust in structure and function in response to experience (Draganski et al., 2004; May, 2011). This new understanding emerged, in part, due to studies demonstrating plasticity modulated by new experiences. For example, London taxi drivers demonstrated greater hippocampal structure associated with spatial navigation learning and updating (Maguire et al., 2000; Maguire, Woollett, & Spiers, 2006; Woollett & Maguire, 2011). Other investigations demonstrated that learned activities, such as juggling, identifying new colours, mirror reading, deciphering Morse Code, and learning new facts, have been robustly associated with increased gray and white matter volumes in areas which support the performance of these tasks (Boyke, Driemeyer, Gaser, Buchel, & May, 2008; Draganski et al., 2004, 2006; Driemeyer, Boyke, Gaser, Büchel, & May, 2008; Ilg et al., 2008; Schmidt-Wilcke, Rosengarth, Luerding, Bogdahn, & Greenlee, 2010; Scholz, Klein, Behrens, & Johansen-Berg, 2009). Although current technologies limit investigators' ability to clearly identify the underlying biological mechanisms associated with neuroanatomical change arising from learning or training in humans, animal models suggest mechanisms such as angiogenesis, increased dendritic branching and synaptogenesis associated with enhanced growth factor gene expression, cell proliferation, and hippocampal neurogenesis (van Praag, Kempermann, & Gage, 2000; Zatorre, Fields, & Johansen-Berg, 2012).

Activity congruent structural brain changes in gray matter have also been demonstrated after working memory training in healthy young adults (Metzler-Baddeley, Caeyenberghs, Foley, & Jones, 2016; Takeuchi et al., 2011, 2010). Metzler-Baddeley and colleagues (2016) randomized participants into either an adaptive or a non-adaptive working memory training intervention using a series of visual and auditory span tasks. Findings in the adaptive training group suggested increased gray matter volumes in the right caudal middle frontal, pars opercularis, and pars triangularis, as well as the left pallidum (although changes were not robust to correction for multiple comparison). Paradoxically, Takeuchi and colleagues (2011) identified decreased gray matter volumes in the bilateral dorsolateral prefrontal cortex, right inferior parietal lobule, left paracentral lobule, and left superior temporal gyrus associated with training on a mental arithmetic working memory task. The authors speculated that gray matter volumes demonstrate a nonlinear function related to working memory training, with volume decreases due to selective elimination of synapses, followed by a reconstruction of structure resulting in volume increases over time (Takeuchi et al., 2011). Considering the brevity of the training period (i.e., 5-days), post-training images may have been obtained prior to reconstruction.

Gray matter volume is an important and frequently studied metric in training studies (e.g., Boyke et al., 2008; Draganski et al., 2004, 2006; Driemeyer et al., 2008; Ilg et al., 2008; Schmidt-Wilcke et al., 2010; Scholz et al., 2009) including studies specific to working memory training (e.g., Metzler-Baddeley et al., 2016; Takeuchi et al., 2010; Takeuchi et al., 2011). Gray matter volume provides a gross estimate of neuronal matter (e.g., cell bodies, dendritic branches) available for use within a particular region, and is robustly associated with cognitive proficiency across the life-span (Andreasen et al., 1993; Frangou, Chitins, & Williams, 2004; Reiss, Abrams, Singer, Ross, & Denckla, 1996; Zimmerman et al., 2006). In their investigation of healthy adults, Andreasen and colleagues, (1993) found that greater gray matter volume was associated with increased performance on numerous individual cognitive tasks and on overall intelligence estimates. Given the cross-sectional study design and correlational nature of results, their findings could be interpreted to suggest that pre-existing gray matter volumes (e.g., genetically derived) facilitate expression of cognitive ability (e.g., performance on cognitive tasks), or alternatively, that pre-existing ability and opportunity to engage in cognitive activities leads to increased gray matter volumes. Although the true relationship is likely much more complex, a four-year study in healthy adults training as London taxi drivers robustly demonstrated that training, rather than pre-existing abilities or brain structures, was associated with structural brain changes in requisite brain areas (Woollett & Maguire, 2011). The specifics of change notwithstanding, the strong association between gray matter volumes and cognition is clearly demonstrated by studies comparing neuroanatomical and cognitive measures (Andreasen et al., 1993; Frangou et al., 2004; Reiss et al., 1996).

Although gray matter volume is an important metric to investigate in training studies, studying volume alone is not sufficient. Mathematically, volume is a function of surface area and thickness. However, neuroanatomical surface area and thickness are two relatively independent metrics influenced by different factors; therefore, studying volume alone may obscure important detail regarding the effects of training (Colom et al., 2016). Like volume, cortical thickness is a common metric of gray matter integrity, with thinning gray matter strongly associated with cognitive dysfunction and preserved gray matter indicative of cognitive health and performance (Dickerson et al., 2008). Furthermore, some researchers suggest that cortical thickness, more so than volume, relates to intellectual ability (Luders, Narr, Thompson, & Toga, 2009; Yuan, Voelkle, & Raz, 2018). Relatedly, surface area is associated with cognitive performance and overall intelligence (Schnack et al., 2014). Surface area measurements utilize information regarding columnar organization, and the principle that neurons within the same column function similarly (Rakic, Ayoub, Breunig, & Dominguez, 2009). Cortical reorganization occurs in response to input or activity (Elbert & Rockstroh, 2004); therefore, cortical surface area, in addition to cortical thickness, is a reasonable target of investigation in working memory training studies. Finally, white matter volume is suggestive of structural integrity which has a direct association with cognitive performance (Johansen-Berg, 2010), and has been shown to increase in response to working memory training (Salminen, Mårtensson, Schubert, & Kühn, 2016; Takeuchi et al., 2010).

A series of recent studies that exclusively used the n-back working memory task revealed training related structural changes in frontal and parietal areas which varied based on the metric utilized (Colom et al., 2016; Román et al., 2016). Specifically, Colom and colleagues (2016) identified increased gray matter volumes in the left posterior cingulate cortex (*d* = 1.15), right cerebellum (*d* = 1.09), and right temporal lobe (*d* = 1.09) in a dual n-back working memory training group relative to passive controls. In a follow-up study, Román and colleagues (2016) identified significant differences in cortical thickness in the right ventral frontal and right middle temporal cortex of trained participants, and significant differences in cortical surface area in the right pars opercularis and right posterotermporal cortex relative to controls. Such findings demonstrate that volume, thickness, and surface area are all useful metrics for identifying neural integrity and change associated with training, and each form of structural integrity can provide unique information regarding the impact of working memory training on brain morphology.

Furthermore, regarding white matter, Takeuchi and colleagues (2010) investigated healthy young adults who completed eight weeks of working memory training consisting of a visuospatial span task, an n-back operation task, and a dual n-back task. Training dosage was positively correlated with increased fractional anisotropy (a measure of white matter fiber tract integrity suggesting increased myelination) in regions implicated in working memory ability, specifically, the intraparietal sulcus and anterior corpus callosum, although conclusions were limited because the study lacked a control group (Takeuchi et al., 2010). However, Salminen and colleagues (2016) included both an active and passive control group, and identified significant increases in fractional anisotropy in the n-back group relative to both control groups, particularly in fronto-parietal, temporo-parietal, and temporo-frontal pathways and the corpus collosum. Although these are the only two studies to investigate white matter changes in response to working memory training, they support the effects of working memory training on white matter structure in healthy adults. Collectively, preliminary findings point toward positive effects of working memory training on gray and white matter structure in healthy adults.

1.6 N-back Working Memory Training and Transfer in Healthy Adults

Our research team recently completed a series of investigations grounded within the P-FIT theory of intelligence as it relates to working memory training and transfer to fluid intelligence (Clark, Lawlor-Savage, & Goghari, 2017a, 2017b, 2017c). Following the recommendations of others (e.g., Buschkuehl & Jaeggi, 2010; Melby-Lervåg & Hulme, 2013) we designed a study that included a homogenous sample of healthy community dwelling adults who we randomly assigned into an n-back working memory training program (experimental condition), an active control condition, or an untrained (i.e., passive) no-contact control condition. We included numerous measures of working memory and fluid intelligence outcomes of interest, and group sample size exceeded the recommended minimum of 20 participants per group (Melby-Lervåg et al., 2016; Simmons et al., 2011). To compliment and extend previous behavioural findings, all training participants underwent structural and functional magnetic resonance imaging (MRI and fMRI respectively) before and after completing their training program. By homogenizing our sample (healthy adults) and training program (n-back working memory training) while expanding our cognitive outcome tasks and including neuroimaging, we hoped to provide clear and objective evidence for or against the effects of working memory training.

The first study (Clark et al., 2017a) examined fMRI patterns for working memory and fluid intelligence tasks prior to training. The intention of this study was to demonstrate that these two key cognitive constructs shared activations in frontal and parietal areas. Although congruent activations during the performance of working memory and fluid intelligence tasks was demonstrated in previous studies, ours was the first to investigate associations between these two constructs in the same sample. Furthermore, the study design allowed for a relatively large group (N = 63) of participants relative to most neuroimaging investigations. Findings supported the hypothesis of overlapping areas of activation, commensurate with increased task difficulty, in expected frontal and parietal areas, specifically, dorsolateral, ventrolateral, rostral prefrontal, premotor, and posterior parietal cortices (Clark et al., 2017a).

The next study (Clark et al., 2017c) compared cognitive performance on working memory and fluid intelligence tasks before and after either n-back training comprised of visual, auditory, and dual n-back training tasks (experimental condition), processing speed training (active control condition), or no training (passive control condition). Contrary to hypotheses and despite significant improvement on the training task itself, the n-back group did not perform better on cognitive outcome tasks after training, relative to either the active or passive control conditions. Null findings were consistent with a similarly designed study that compared healthy adults (age 30 to 60 years) on the same key cognitive outcomes after exclusively training with the dual n-back task, but that lacked a passive control group to discriminate practice effects from training effects shared between the active control and experimental conditions (Lawlor-Savage & Goghari, 2016). Although literature published before the conceptualization of our current group of studies suggested we could expect large near transfer and small but significant far transfer effects (Melby-Lervåg & Hulme, 2013), and more recent meta-analyses support that expectation (Au et al., 2015, 2016), other recent meta-analyses were consistent with our null behavioural findings (Dougherty et al., 2016; Melby-Lervåg & Hulme, 2016).

The third study (Clark et al., 2017b) examined changes in functional brain activity before and after n-back working memory training or processing speed (active control condition) training.

Findings indicated that, relative to active controls, activation patterns during performance of a dual n-back task decreased in working memory trainees after six weeks of n-back training. This outcome suggested increased efficiency within associated networks due to repeated working memory task practice. However, no changes were detected in activation patterns between the two training groups during the performance of a fluid intelligence task. Functional activation data, combined with cognitive task data, suggest that in these studies, working memory may have had an impact on neural functioning in the absence of far transfer to cognitive outcomes.

Buschkuehl, Jaeggi, & Jonides (2012) described three possible ways to interpret fMRI data related to training, independent of whether cognitive task performance corresponds to fMRI activations. First, increased activity in a particular region of interest may be associated with improved performance on a cognitive task due to increased cortical representation or more robust neural responding as more neural resources are allocated to that area. Conversely, increased efficiency of neural processes related to a task may result in decreased fMRI activity with increased cognitive task performance. Finally, activity may both increase and decrease across different neural regions associated with the training task (i.e., activation redistribution) or new neural areas may be recruited and thus activate in response to a particular task (i.e., activation reorganization). Furthermore, previous investigations have identified neuroanatomical changes associated with working memory training in the absence of changed cognitive task score (Engvig et al., 2010). Hence, while it is useful to refer to fMRI activation patterns associated with cognitive training in order to guide hypotheses regarding regions expected to change following training, fMRI studies fail to provide concrete information pertinent to brain reserve.

The first three studies published in our overall investigation focused on potential cognitive enhancements and functional activations associated with working memory training. Although the results of these investigations do not support the presence of near or far transfer to cognitive task performance associated with working memory training, important areas of investigation remain unaddressed, specifically, neuroanatomical changes associated with training. Before thoroughly concluding that in this sample, working memory training was not effective at influencing the brain, morphometry must be examined in order to confirm, or disconfirm, the absence of training dependent neural change. Therefore, the data I present in this document are intended to complement and extend the findings of our previous investigations by examining the neuroanatomical correlates of working memory training and transfer.

1.7. Rationale and Hypotheses

The purpose of the present study is to investigate alterations in neuroanatomy associated with working memory training and transfer to fluid intelligence. Specifically, this study asked (1) for healthy adults, does working memory training alter gray matter volume, thickness, or surface area in a priori regions of interest implicated in working memory and fluid intelligence? Regions of interest were consistent with the P-FIT theory of intelligence, and neuroimaging investigations of working memory and the effects of working memory training. Specifically, regions of interest within the frontal lobe were the superior, rostral middle, and caudal middle frontal gyri, the pars opercularis, triangularis, and orbitalis, the precentral gyrus, and rostral and caudal anterior divisions of the cingulate cortex. Parietal regions included the postcentral and supramarginal gyri, and the superior and inferior parietal cortex, as well as the adjacent insular cortex. Regarding subcortical regions of interest, I looked for volume increases in the caudate, putamen, pallidum, thalamus, and hippocampus. In addition to examining specific regions of interest, this study also asked (2) does working memory training influence overall gray and white matter volumes? Finally, given known associations among age, estimated intelligence (i.e., IQ), and structural metrics, I investigated the extent to which baseline individual factors influence change by asking (3) are lower baseline structural metrics, lower participant age, or lower baseline estimated intelligence scores associated with greater structural change after working memory training?

When this study was initially proposed, there was sufficient evidence to hypothesize that in this sample, working memory training would be associated with enhanced cognitive task performance. Specifically, a comprehensive meta-analysis of available literature revealed improved working memory (near transfer, d = 0.79) and fluid intelligence (far transfer; d = 0.34) task performance after n-back working memory training (Melby-Lervåg & Hulme, 2013). Based on the assumption of detecting changes in near and far transfer after training, I initially planned to investigate associations among changes in cognitive performance and structural change. Specifically, I wanted to ask if lower baseline structural metrics were associated with greater levels of cognitive change after training, or if change on cognitive measures correlated with structural change. However, recently published analyses of the cognitive data collected as part of the larger collection of our studies failed to support changes in cognition associated with working memory training (Clark et al., 2017c), rendering further examination of structural and cognitive correlates irrelevant. Nevertheless, in this paper I included analyses of cognitive task scores for the two training groups for two reasons: (1) to provide more detail (e.g., group, time, and interaction effects) than was previously published (Clark et al., 2017c), and (2) because, although the sample that underwent neuroimaging was the same, there were data acquisition issues with one participant which slightly altered the composition of the samples for which functional versus structural analyses were completed. Hence, the cognitive task data I present are specific to the sample that underwent structural analysis.

The primary hypothesis of this study is multicomponent. First, I expect the n-back working memory training group, relative to the active control group, to exhibit increased gray matter volume, surface area, and thickness measures in regions of interest within the frontal, parietal, and insular cortex after training. Similarly, I predict that the working memory training group will demonstrate greater gray matter volumes in specific subcortical regions of interest: caudate, putamen, pallidum, thalamus, and hippocampus. Finally, I anticipate greater overall gray and white matter volumes in working memory trainees relative to active control participants.

The secondary hypotheses relate to individual differences associated with neuroanatomical change after working memory training. I predict that overall baseline gray and white matter volumes will be negatively correlated with training related change. Specifically, I anticipate that lower baseline brain matter will be associated with greater brain matter change after training in the working memory training group. Further, I anticipate that lower participant age will be associated with greater change. Finally, given that overall intelligence is associated with neural structure, I expect that lower estimated baseline intelligence will be associated with greater structural change after training.

Chapter Two: Methods

2.1 Participants and Recruitment

Participants for this study were recruited as part of a larger overall study investigating neural correlates of working memory training, components of which have already been published (Clark et al., 2017a, 2017b, 2017c). Adults aged 18 to 40 years were recruited from the Calgary community via community postings and word of mouth. Interested individuals were directed to the study website www.braintrainingstudy.ca where they expressed their interest in the study by completing a brief screening questionnaire that was used to assess for inclusion and exclusion criteria (see Appendix A for screening questions). Exclusion criteria were visual or auditory impairment (not corrected by glasses, contacts, or hearing aids); psychotropic drug use in the past three months (including prescription); lifetime history of brain trauma; neurological or psychiatric illness; or current pathology associated with cognitive impairment including renal, respiratory, cardiac, or hepatic disease, uncontrolled diabetes, or cancer not in remission more than two years (Crook et al., 1986). Those who regularly (e.g., more than twice per week) used an n-back or processing speed training product in the previous six months were also excluded. Additionally, those unable to tolerate scanning protocols (e.g., due to claustrophobia, metal in body, etc.) were excluded. Study inclusion required that participants were right handed, healthy, aged 18-40 years, had access to a desktop or laptop computer with high-speed internet, could devote 30 minutes to training five days per week, and could attend in-person cognitive assessments and scanning appointments. A flow chart of the study design is presented in Figure 1.

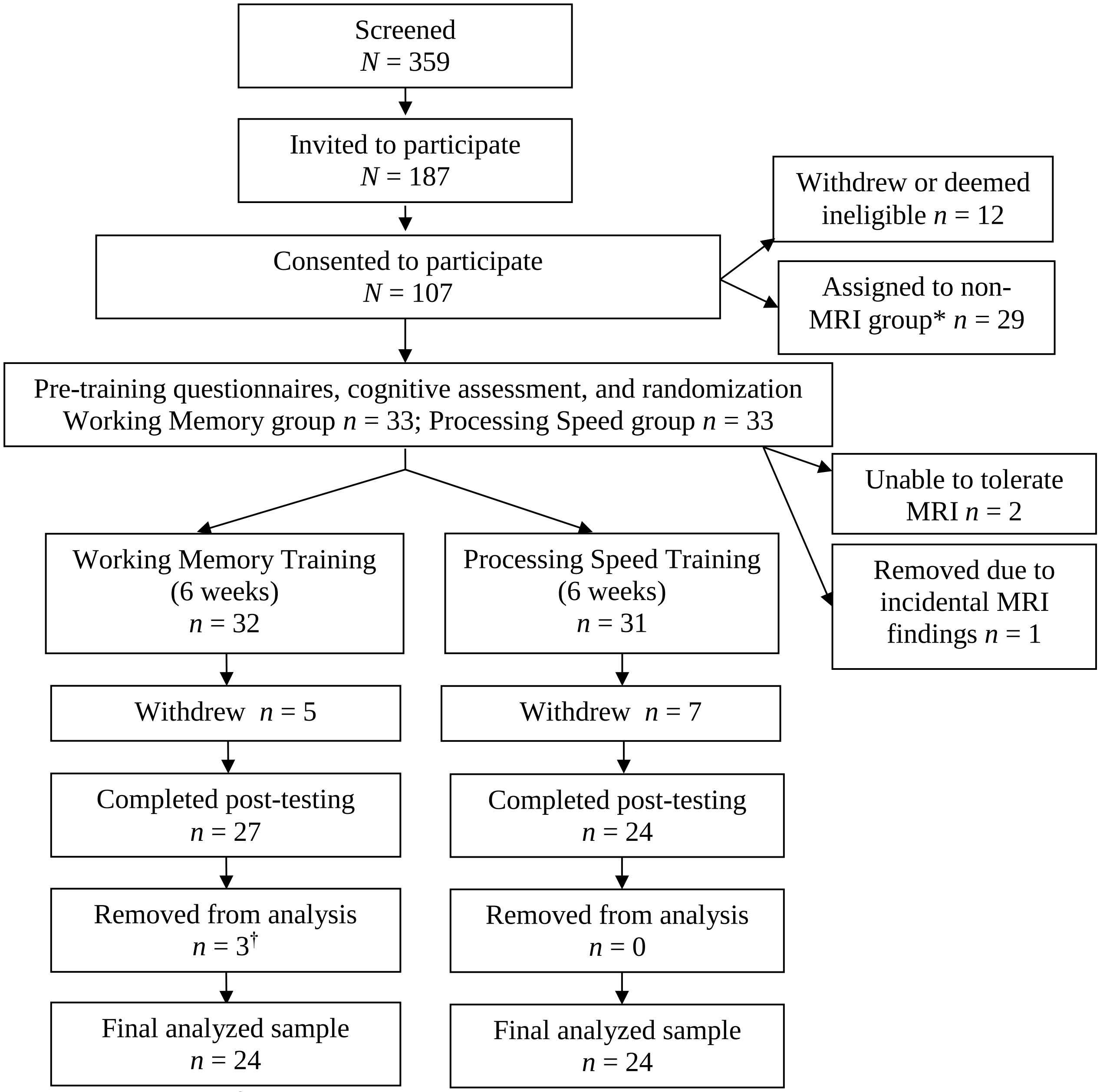

Figure 1. Flow chart of study design

*Recruitment and data collection was shared with a study which utilized a no-MRI condition.

†Participants removed from analysis due to poor compliance (i.e., training contamination and/or low dosage; *n*=2), or data acquisition issues (*n*=1).

The age range was chosen for multiple reasons. First, most studies have been conducted in young adults (e.g., under 30) so by studying the 18 to 40 year age range we were able to compare our results with existing findings, while still allowing for our findings to extend to a broader age range, particularly, those approaching middle age. Second, numerous studies have suggested that many cognitive abilities, and aspects of brain structure including cortical thickness, begin to decline in the third decade of life (Salat et al., 2004; Salthouse, 2009). On average, the sample in this study was expected to fall around their highest lifetime level of memory and reasoning ability (Salthouse, 2009) and thus any cognitive or neural changes associated with training would be considered a strong marker of the potential to strengthen these abilities and their neural correlates. Finally, observation of advertisements of commercial cognitive training programs indicate that these programs are marketed to individuals within our study's age range; therefore, results of this study, whether supporting or countering claims of the effectiveness of training, would result in useful information for society in general.

The sample size was determined in consideration of our overall project investigating neural correlates of working memory training. An initial power calculation was based on a meta-analysis of n-back training studies with cognitive task outcomes, which revealed a large effect of n-back training on near transfer (d = 0.79) and a small to moderate effect on fluid intelligence (d = 0.34; Melby-Lervåg & Hulme, 2013). Power was estimated using G*Power, a freely available software program composed of algorithms for calculating statistical power for a variety of families of tests, including F-tests utilizing repeated measures assessing within and between group interactions (Erdfelder, Faul, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007). To achieve a power of 0.8 with $\alpha = .05$ in a 3 (group) x 2 (time) repeated measures Analysis of Variance (RM-ANOVA) design based on an anticipated Cohen's d = 0.34 (based on the expected behavioural far transfer effect of primary interest), a total of 87 participants were suggested. However, the same calculation using a 2 (group) x 2 (time) RM-ANOVA design, as was used for the data presented in this dissertation, returned a total of 70 participants, or 35 per group. In terms of determining power for the neuroimaging aspect of the study, when this study was initiated there were no similar studies from which to draw expected effect sizes for structural changes associated with dual n-back training. However, feasibility and budgetary constraints with the neuroimaging aspect of the study limited the total number of participants we could scan to 60, allowing for 30 per training group less an expected attrition rate of 25% (consistent with a recent study conducted within the same community; Savage, 2013). Given available resources, expected attrition rates, and the recommended minimum of 20 participants per group (Melby-Lervåg et al., 2016; Simmons et al., 2011), we aimed to recruit 30 individuals per group and expected to analyze data from 20 - 25 participants per group.

All participants were compensated $20 per appointment attended (i.e., cognitive testing session, MRI session) totalling $80 for the four appointments attended by participants who completed all components of the study. Written consent was obtained from all participants, and this study was approved by the University of Calgary's Conjoint Health Research Ethics Board (CHREB).

2.2 Procedure

Potential participants who met inclusion criteria were provided with details about what the study entailed, and were invited to participate. Individuals who accepted the invitation were emailed links to an online consent form and comprehensive questionnaires assessing demographic and individual characteristics: age, gender, ethnicity, language, martial status, employment status, education, income, health history including current and past conditions and drug use, cognitive activities, cognitive training history, handedness, and height and weight. A copy of the questions can be viewed in Appendix B. Responses were reviewed to ensure that the individual met criteria for the study, with ineligible participants removed from the study at this point. Eligible and consenting participants were then scheduled to attend a baseline MRI scanning session followed by an in-person cognitive testing session within five days of the MRI scan.

Participants were randomized to one of two training groups: a working memory training group (i.e., the experimental condition), or a processing speed training group (the active control condition). Processing speed training was chosen as a control condition because, although training of processing speed has robust effects on improving performance on measures of processing speed, such training does not extend to improvements in other cognitive domains (Takeuchi & Kawashima, 2012). Specifically, across numerous investigations, processing speed training has improved performance on the trained task (i.e., practice effects) but not untrained tasks assessing fluid intelligence, working memory, memory encoding and recall, daily problem solving, and speeded tasks which differed from the trained task (Ball et al., 2002; Edwards et al., 2005; Edwards et al., 2002; Lawlor-Savage & Goghari, 2016; Mackey, Hill, Stone, & Bunge, 2011; Takeuchi & Kawashima, 2012; Vance et al., 2007; Wadley et al., 2006; Willis et al., 2006). Although no cognitive task is pure (e.g., an individual with a fast thinking speed is able to process more information in a discrete unit of time), the processing speed task required only minimal demand on working memory (i.e., identifying a shape and quickly deciding if it was the same as the previous shape) relative to a complex working memory task (i.e., holding multiple pieces of visual and auditory information in mind while updating that information and responding to new information) which elicits a high working memory load. Therefore, including processing speed training as an active control group allowed for all aspects of the investigation to be held constant between the two groups (e.g., interaction with researchers, engagement with the training platform, and levels of effort and motivation) except the specific training content (i.e., the independent variable).

Randomization to training group was determined by random number generator and managed by a research assistant who was assisting with the preparation of study materials but was not otherwise involved in the study and did not have contact with study participants. Randomization to training group was revealed to both the participant and the test administrator at the end of the first cognitive testing appointment, when the participant opened a sealed envelope that revealed instructions on how to access training. Therefore, the tester working with the participant was blind to training group during the baseline MRI and cognitive testing sessions. To maintain blinding, post-training sessions were assigned to a different tester, although due to scheduling, this arrangement was not always possible. Therefore, in a strict sense, investigator blinding was not truly preserved throughout the study. However, participants were asked to not discuss their training experience until after testing was complete and investigators generally reported not being able to identify to which training group the participant belonged. Furthermore, testing itself did not reveal randomization. Although the decision to introduce a different tester at the post-training session had the potential to introduce subtle inter-examiner differences, all testers were similarly trained and supervised, including observation of their interactions with participants, and the tests themselves were standardized.

As part of the baseline MRI appointment, participants were reminded of the study requirements, completed written consent, and completed a questionnaire package. Details regarding these questionnaires and the MRI data acquisition procedure are described in subsequent sections.

During the baseline and post-training cognitive testing sessions, all participants completed the same cognitive measures (described below). Cognitive tasks were administered in a rolling order (rather than randomized order) to account for fatigue effects. Given our limited sample size, we were unable to statistically measure for task order effects therefore decided to balance fatigue effects. Furthermore, in this experiment, waning cognitive effort due to fatigue was expected to have more influence on performance than the order in which the tests were presented. Test order was repeated at the post-training cognitive testing session. All cognitive testing was completed by myself and another PhD candidate with training and experience conducting neuropsychological assessments, or one of two upper year undergraduate or post-undergraduate research assistants who I trained and supervised.

All training programs were accessed by participants through the website www.lumosity.com. Participants were directed to their assigned training programs based on individual log-in credentials associated with either working memory or processing speed training tasks. With the exception of the specific games, all aspects of the training program (e.g., training platform, including online access through the Lumosity website, visual attractiveness of the game, and training length) were the same between the working memory and processing speed training programs.

2.3 Training Programs

2.3.1 N-back Training

The n-back training program (experimental training) was composed of three games accessed through an online training platform (Lumos Labs Inc., 2009): Memory Lane, Memory Match, and Memory Match Overload. Memory Lane is a visually embellished version of the adaptive dual n-back task used in previous studies to demonstrate near and far transfer effects of training in healthy young adults (Jaeggi et al., 2010, 2011, 2014; Jaeggi, Buschkuehl, et al., 2008; Rudebeck et al., 2012; Schweizer et al., 2011; Stephenson & Halpern, 2013). As such, Memory Lane is a cognitively complex dual n-back task that requires simultaneous use of auditory and visual processes at increasing cognitive loads (i.e., difficulty) based on performance. On the screen, the participant saw an image of a building with four to nine windows (depending on game difficulty level). A silhouette appeared in one of the windows, and a letter was spoken, comprising one trial of the dual n-back task. Then, a new building appeared with a silhouette and spoken letter, representing the second trial, and so on for 15 trials. Participants were instructed to respond via left arrow keyboard press when the silhouette position matched n-back (i.e., 1-back as one building prior, 2-back as two buildings prior, etc.) and right arrow when the auditory stimulus matched n-back. After each round, game difficulty was adjusted based on individual performance. Difficulty increased for task performance meeting or exceeding 80% accuracy, and decreased for performance below 80% accuracy. Cognitive load was adjusted across one of three possible dimensions. Specifically, the grid of windows on the building varied (i.e., 2x2, 2x3, and 3x3), the target n (number of buildings previous) ranged from 1-back to 10-back, and the stimuli modality changed from single n-back presented visually only, to dual n-back presented both visually and auditorily. Therefore, the task fit within the mismatch model of neural and cognitive plasticity (Lövdén et al., 2010) because a gap was continually created between a participant's functioning (i.e., ≥80% accuracy) and the cognitive demands of the task (i.e., the next level of difficulty). After each set of three rounds, a game score was generated based on response accuracy.

Memory Match and Memory Match Overload were both visual n-back tasks that were included in the training package to allow for a greater variety of games, and an increased likelihood of engagement, relative to the dual n-back task in isolation. Memory Match was a visual 2-back task where participants saw a sequence of ordered shapes moving across the screen from right to left, with one shape per trial. Participants were to identify when the target shape (i.e., current trial, in the right-most position) matched the shape presented 2-trials previously (occupying the left-most position). The task, therefore, required participants to continually encode, retain, and update the sequence of information to produce a correct response. For the first few trials, all shapes were visible moving across the screen from right to left; however, after successful completion of a few trials all shapes other than the target shape became invisible. If performance accuracy fell below 80%, visibility of the shapes returned until performance accuracy returned to 80% or above. Memory Match Overload was a visually similar, but slightly more difficult, 3-back version of Memory Match. Accuracy scores were generated after each Memory Match and Memory Match Overload game. Screen shots of the instructions for each game are presented in Figure 2.

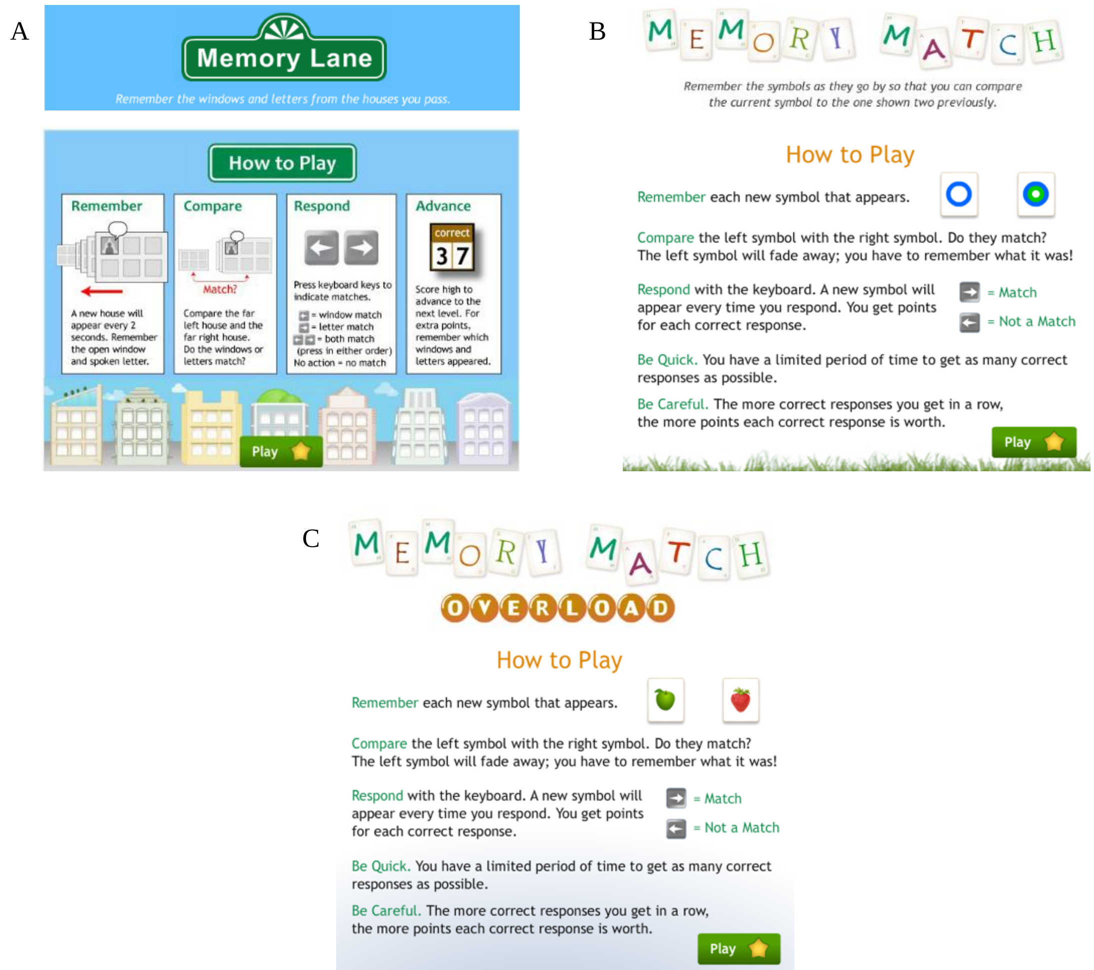

Figure 2: Screen captures of instructions given to participants for how to play and earn points on A) Memory Lane, B) Memory Match, and C) Memory Match Overload working memory training games.

Game durations were 180 seconds (consisting of three 60-second rounds) for each dual n-back Memory Lane game, 45 seconds for each 2-back Memory Match game, and 45 seconds for each 3-back Memory Match Overload game. Each training session consisted of six Memory Match games, five Memory Match Overload games, and five Memory Lane games for a total training session time of approximately 24.5 minutes.

2.3.2 Processing Speed Training

The processing speed training program, included as the active control training condition, was composed of three games accessed through the same online training platform (Lumos Labs Inc., 2009): Speed Match, Speed Match Overdrive, and Spatial Speed Match. In each of the three processing speed training games, participants attended to and quickly responded based on features of the stimuli, specifically, determining whether a presented symbol matched a symbol presented immediately before. Speed Match and Spatial Speed Match were visual 1-back tasks where participants saw a sequence of shapes and had to determine if the current shape exactly matched the shape presented immediately prior. Unlike the Memory Match tasks, shapes remained visible during the Speed Match trials. Speed Match Overdrive was similar but included a partial match response option. Specifically, participants had to quickly decide if the second shape matched the first shape partially (down arrow button press), exactly (right arrow button press), or not at all (left arrow button press). Although these games were expected to recruit basic working memory processes, such as immediate attention and short-term maintenance of small amounts of information, the focus of these training games were on reaction speed rather than maintenance, manipulation, and complex processing of multiple pieces of information in working memory. The three speed games (Speed Match, Speed Match Overdrive, Spatial Speed Match) lasted 45 seconds each and were presented 11 times per training session for a total of approximately 24.75 minutes of training per session. Screen-captures of the speed-based training games are presented in Figure 3.

Figure 3: Screen captures of instructions given to participants for how to play and earn points on A) Speed Match, B) Spatial Speed Match, and C) Speed Match Overdrive processing speed training games.

2.4 Individual Characteristics and Cognitive Tasks

2.4.1 Baseline Individual Characteristics.

All participants completed an online form assessing demographics (i.e., age, gender, ethnicity, marital status, education, employment status, income) and a variety of characteristics that have the potential to influence cognitive performance. These individual characteristics data were collected in order to identify, and if necessary control for, baseline differences between the experimental and the control group. Mood was measured with the Profile of Mood States-Short Form (POMS-SF; Curran, Andrykowski, & Studts, 1995), sleep with the Pittsburgh Sleep Quality Index (PSQI; Buysse et al., 1989), physical activity with the International Physical Activity Questionnaire – Last 7 Day, Short Form, Self-Administered (IPAQ-S7S; Craig et al., 2003), and personality with the HEXACO-60 (Ashton & Lee, 2009). All measures have demonstrated good psychometric properties in healthy adults. An additional questionnaire was created for use in this study to assess participant motivation and training related expectations given that differences in motivation among groups can artificially influence training effects (Boot, Simons, Stothart, & Stutts, 2013; Foroughi, Monfort, Paczynski, McKnight, & Greenwood, 2016). Questionnaires assessing demographic and individual characteristics are in Appendix B and C.

2.4.2 Cognitive Outcome Measures

2.4.2.1 Wechsler Adult Intelligence Scale

The Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV; Wechsler, 2008) is a highly valid and reliable estimate of overall intelligence as well as particular facets of intelligence, specifically, verbal comprehension, perceptual reasoning, working memory, and processing speed. Eight of the 10 core subtests of the WAIS-IV were administered: Vocabulary, Similarities, Block Design, Matrix Reasoning, Digit Span, Arithmetic, Symbol Search, and Coding. The two remaining subtests, Information and Visual Puzzles, were excluded to save time and participant burden since the WAIS-IV full scale and composite index scores can be calculated using only the eight administered tests (Wechsler, 2008).

Pre- and post-assessment measures were administered less than two months apart; therefore, the Vocabulary, Similarities, Block Design, Matrix Reasoning, and Arithmetic subtests were split into two forms, each composed of even or odd numbered items in order to limit practice effects. Odd forms were administered at baseline and even forms were administered after training. The structure of these subtests allowed for them to be easily split into parallel forms; however, the remaining three subtests (Digit Span, Symbol Search, and Coding) could not be easily administered as split forms and were therefore administered in full at baseline and again at post-training testing sessions. Split-half reliability estimates of the WAIS-IV range from .71-.96 for individual subtests, .87-.98 for Index scores, and .97-.98 overall (Wechsler, 2008). Test-retest reliability coefficients were calculated using scores from a no-contact control group collected as part of the overall study and are presented for tests analyzed in the present study: Full-Scale IQ $r = .91$, Block Design $r = .66$, Matrix Reasoning $r = .34$, Digit Span, $r = .82$, Arithmetic $r = .69$, Symbol Search $r = .59$, Coding $r = .49$ (Clark, 2017). Of note is that reliability estimates used in our studies (Clark, 2017) are superior to those reported in the WAIS-IV manual (Wechsler, 2008) because in our studies, the forms were split (i.e., different individual items were administered), and the test-retest period was greater (i.e., approximately 6-weeks compared to 2-weeks). Hence, I am confident that administering the measure in a non-standardized manner by utilizing split-forms was an appropriate trade-off to ensure potential gains were not due to familiarity with particular test questions. Due to the split-half procedure, raw rather than scaled scores were utilized as outcome measures. Full-scale intelligent quotient scores at baseline were estimated by doubling raw baseline scores, and standardized discontinue rules were divided by two, and rounded up.

In terms of individual WAIS-IV subtests, the Digit Span and Arithmetic subtests were categorized as measures of working memory, Block Design and Matrix Reasoning as measures of perceptual reasoning, and Coding and Symbol Search as measures of processing speed. The Vocabulary and Similarities subtests were administered as part of the overall study that addressed a different theoretical question and are therefore not included in the data presented for this study, beyond their contribution to overall estimates of intelligence.

Regarding working memory, the Digit Span subtest is particularly useful in measuring components of working memory as the task can be broken down into auditory attention span (i.e., digit span forward) and auditory manipulation (i.e., digit span backward and sequencing), both which correspond to the phonological loop of Baddely's working memory model (Baddeley, 2017; Baddeley, 2003; Baddeley & Hitch, 1974; Baddeley & Logie, 1999). Participants listened to a group of numbers presented one second apart, and repeated the numbers either as heard (digit span forward), in reverse (digit span backward), or in order from lowest to highest (digit span sequencing). The Arithmetic task required the participant to attend to and temporarily store auditory input and bring up rules of mathematical operation from long term stores in order to produce a response; therefore, the arithmetic task invokes the phonological loop as well as higher order executive processes of working memory. Although the phonological loop is of focus in the Digit Span and Arithmetic tasks, no task can perfectly differentiate individual aspects of working memory. For example, visually inclined individuals may also make use of their visuospatial sketchpad to recall or sequence digits, or to answer verbal math questions. Therefore, this study was not able to break working memory processes into discrete components based on Baddeley's model.

Regarding the processing speed tasks, participants were required to quickly respond to visual stimuli by copying symbols on paper or identifying symbols that match other symbols. The WAIS-IV Processing Speed Index is a reliable measure of processing speed capability, and the Symbol Search subtest has been described as one of the most pure processing speed measures (Wechsler, 2008). Finally, non-verbal reasoning was measured with the Block Design and Matrix Reasoning Tasks, which collectively comprise the perceptual reasoning measure on the WAIS-IV (Wechsler, 2008). For the Block Design tasks, participants reproduced a two dimensional image using blocks, and for the Matrix task, participants chose from a group of eight options an option that would correctly complete a sequence or puzzle.

2.4.2.2 Automated Operation Span

The Automated Operation Span (Aospan) task (Unsworth, Heitz, Schrock, & Engle, 2005) is a complex computerized working memory task wherein the participant quickly solved mathematical operations while remembering a sequence of letters. Specifically, the Aospan task required the participant to respond to a math problem, then disregard that information while remembering a letter, then attend to a new math problem while maintaining the previous letter in memory and learning a new letter. Therefore, the task recruited numerous aspects of working memory, including maintenance and updating. Outcome scores on this task include an absolute score representing the sum of all perfectly recalled letter sets, and a 'total number correct' score representing the number of individual letters recalled in correct serial position, regardless of whether all letters within a set were recalled. The total score, rather than the absolute score, was used in this study because the total score is, theoretically, a more sensitive estimate of change. The

Aospan task has been used as an outcome of working memory in cognitive training studies with similar designs as the present study (e.g., Jaeggi et al., 2010; Lawlor-Savage & Goghari, 2016; Thompson et al., 2013) and was included in this study as a measure of complex working memory representing near transfer.

The Aospan task has good psychometric properties, including test-retest reliability of .83 over a median time span of 6 days (M_{days} = 13; range = 172), and is highly correlated with other operation span tasks (Unsworth et al., 2005) and other measures of working memory (Conway et al., 2002). The test-retest reliability coefficient based on the overall study's no-contact control group was r = .82 (Clark, 2017).

2.4.2.3 Spatial Delay Response Task

Working memory was also measured with the Spatial Delay Response Task (SDRT; Glahn et al., 2002). This task assesses both memory span and processing ability for increasing amounts of spatial information, and has been associated with neural activation in working memory areas (Glahn et al., 2002). This task is composed of two conditions. First, participants were briefly presented with an array of 1, 3, 5, or 7 yellow circles on a screen followed by a two second delay. A second, equal numbered, set of circles was presented, and the examinee decided if the second set of circles was in the same position as the first set. This condition relies on spatial maintenance, and the resulting *spatial maintenance* score is derived from participants' accuracy (% correct) at remembering the positions of circles. The second condition is similar, except that after the two second delay, the circles are presented flipped about the horizontal midline. Participants were required to recall the original position of the circles and mentally flip them in memory to determine if the flipped position of the newly presented circles matched the flipped position of the previously presented stimuli. The *maintenance plus manipulation* score was obtained from participants’ ability (% correct) to correctly determine if the second presentation matches the spatially manipulated (i.e., flipped) original presentation. Including this task allowed for a measure of visuospatial maintenance corresponding to Baddeley’s model of working memory (Baddeley, 2003, 2017; Baddeley & Logie, 1999; Baddeley & Hitch, 1974) specifically, the visuospatial sketchpad. However, again, we recognize that participants may have also engaged phonological processes, therefore this task is considered a working memory measure but not specifically a measure of visuospatial processing. The test-retest reliability coefficients based on the overall study’s no-contact control group was $r = .73$ for the more challenging maintenance plus manipulation task, although reliability of the maintenance only task was poor ($r = .02$; Clark, 2017). Both task conditions were included as measures of near transfer to working memory, with the maintenance task representing simple visual span and the manipulation task an estimate of complex working memory.

2.4.2.4 Raven’s Advanced Progressive Matrices

The Raven’s Advanced Progressive Matrices (RAPM; Raven, 1975) is a matrix-based measure of fluid intelligence that highly converges with other measures of fluid intelligence and general intellectual ability (Carpenter, Just, & Shell, 1990; NCS Pearson Inc., 2007). In each RAPM item, the participant looked at a 3 x 3 grid of images with one square missing. The participant chose the best of eight possible images to complete the grid. The first set of images was a practice set of 12 items that quickly progressed from easy to challenging. Points were not awarded for practice items. The second set contained 36 items which progressed from easy to challenging; one point was scored for each correctly answered item.

The RAPM is commonly used as a measure of fluid intelligence in dual n-back working memory training studies (Jaeggi et al., 2010; Jaeggi, Buschkuehl, et al., 2008; Jaušovec &

Jaušovec, 2012; Lawlor-Savage & Goghari, 2016; Salminen et al., 2012; Thompson et al., 2013). Given the high internal consistency of the RAPM (split-half reliability .85; NCS Pearson Inc., 2007), and consistent with previous studies (Jaeggi et al., 2010; Jaeggi, Buschkuehl, et al., 2008; Lawlor-Savage & Goghari, 2016; Salminen et al., 2012), we split this test into parallel odd and even forms and administration time was halved. Total correct responses were used as an outcome measure of fluid intelligence, representing far transfer. The test-retest reliability coefficients from the overall study's no-contact control group (n = 23) was r = .55 (Clark, 2017).

2.4.2.5 Cattell's Culture Fair Test

Cattell's Culture Fair Test (CCFT; Cattell & Cattell, 1959) Scale 3, Forms A and B consist of four subtests measuring fluid intelligence: series, classifications, matrices, and conditions. Participants viewed a series of images and chose the correct image, from four possibilities, to complete the series or puzzle. The CCFT is described as a superior and more specific measure of fluid intelligence compared to measures using only matrices tasks and is appropriate for the general adult population ranging from average to superior intelligence (Colom & García-López, 2003). Forms A and B are designed to be used sequentially therefore form A was administered at baseline and form B after training. When forms A and B are used sequentially, the CCFT has demonstrated good internal (0.85) and test-retest (0.82) reliability, high construct validity (0.92), and adequate criterion validity (0.69) with more general measures of intelligence (Cattell & Cattell, 1959). Test-retest reliability based on our larger study's no-contact control (n = 23) group was r = .69 (Clark, 2017).

2.5 MRI Data Acquisition

Whole-brain digital imaging and communication in medicine (DICOM) images were collected on a 3T General Electric Discovery MR750 system using an 8-channel head coil at the

Seaman Family Magnetic Resonance Research Centre at the University of Calgary. T1-weighted 3D magnetization-prepared rapid acquisition gradient echo (MP-RAGE) anatomical scans were acquired for each participant (256, 1 mm slices, TE = 3.1 ms, TR = 7.4 ms, inversion time (TI) = 650 ms, FOV= 25.6, matrix = 256 × 256).

Structural MRI data was collected as part of a larger protocol which included functional neuroimaging (see Clark et al., 2017a, 2017b) so participants spent approximately 60 to 75 minutes in the scanner alternating between cognitive tasks and rest. Structural scans lasted approximately 10 minutes, for which participants were instructed to relax and remain still.

2.5.1 Individual Image Processing

Images were processed using FreeSurfer image analysis suite, version 5.3.0, a freely available program for cortical reconstruction and volumetric segmentation (FreeSurfer Software Suite). Images were processed using the longitudinal stream (Reuter, Schmansky, Rosas, & Fischl, 2012). Briefly, each set of participant images from each time point underwent a fully automated cortical reconstruction process which includes skull stripping, volumetric labelling, intensity normalization, white matter segmentation surface at last registration, surface extraction, and gyral labeling. The program modeled the surface as a mesh of triangles, and the coordinate points of each triangle’s vertices were visualized as a surface within a three-dimensional space, and used to determine surface area within functional regions. Upon completion of this process, a file of statistics is generated including surface cortical area and thickness (mm), and regional and total gray and white matter volumes (mm3).

After the initial reconstruction process, I visually inspected each image slide for errors in the skull strip, white matter segmentation, white matter intensity, and surface reconstruction automated processing steps. Common errors were scull strip errors in which dura was incorrectly included in the gray matter ribbon. Upon initial review of the images, I applied minor edits to 56 of the 113 reconstructed images based on Freesurfer developer recommendations (D. Greve & B. Fischl, personal communication, September 30, 2015). Offending voxels were removed and the corrected surfaces and volumes were reconstructed. Control points were added to five images due to errors in voxel intensity, which influenced gray and white matter differentiation. After editing was complete, images were submitted to a second reconstruction, revised reconstructions were inspected, and additional minor edits were applied as needed.

Editing was completed with blinding procedures in place. Specifically, prior to the initial reconstruction, file names were replaced with a five-digit number created by a random digit generator. Editing was then completed while blind to participant number, group randomization, longitudinal order (i.e., pre- versus post- training MRI), and order of enrollment and participation (e.g., early versus late in study). Once all editing was complete, files were relabeled with their original names prior to data extraction and analysis.

2.5.2 Longitudinal Processing

As described by Reuter and colleagues (2012), the longitudinal processing stream involved creating a within-subject template representing an average of all time points for each participant. The resulting template was used as a participant specific estimate for the segmentation and reconstruction process of the individual time. As with the individual time points, I manually inspected each base for errors, and applied minor corrections as needed. Cross-sectional images were then processed through the longitudinal stream, which incorporates information from the within-subject template. Advantages of this approach include that all time-points are treated equally (even if only a baseline scan was collected), and variability is reduced because each time-point is registered to a common anatomical base specific to each participant.

The resulting output provided estimates of overall brain volume as well as gray and white matter total volumes, and gray matter thickness and surface area at both time points based on the Desikan-Killiany atlas (Reuter et al., 2012). Cortical thickness was determined by the software as a measure of the distance between the pial surface and the boundary between gray and white matter (Greve, 2011). Surface area of functional regions was quantified based on the points of each vertex in the surface mesh (Greve, 2011). Volume was determined by the program as a function of surface area and thickness (Greve, 2011). White matter volume was calculated by subtracting all non-white matter voxels from overall brain matter volume (Greve, 2011).

2.6 Statistical Analyses

All analyses were conducted using Statistical Package for the Social Sciences (SPSS) version 24. Independent samples *t*-tests and chi-squared analyses were used to explore participant characteristics and baseline differences in demographics, individual characteristics (estimated intelligence, personality, sleep quality, physical activity, mood, expectancy, and motivation), cognitive outcomes, and neuroanatomical features between study completers and non-completers, as well as between training groups. Two-tailed correlations were conducted among all baseline cognitive outcome measures to assess convergent validity of theoretically related cognitive tasks.

2.6.1 Behavioural Analyses

Training characteristics, including amount of time spent training and training-related improvement across individual games (e.g., score increases, or reaction time decreases) were compared using paired samples *t*-tests. Improvement was determined by comparing the average of the first five, and the last five, iterations of each game. This procedure was utilized due to a dramatic spike in performance during the first few attempts at each game, likely reflecting the participant becoming familiar with the platform and games. Averaging these first five attempts provided a more accurate estimate of baseline task performance and reduced the likelihood that effects were artificially inflated due to the task being novel during the first attempt. For consistency, the last five iterations of each game were also averaged.

Cognitive test outcome scores were analyzed within each cognitive domain using separate 2 (group) x 2 (time) RM-ANOVAs to report differences between groups for each cognitive test after training. Alpha levels were adjusted based on the number of tests within each domain (i.e., within each family of ANOVAs). Of note is that the present study compares findings between two training groups, the working memory training experimental group and the processing speed training active control group. A no-contact control group was not included in this analysis as those data were previously published (Clark et al., 2017c). Clark and colleagues (2017c) also published cognitive test data for the working memory training and processing speed training groups; however, slight differences between the two studies merit the presentation of the behavioural data specific to the participants whose structural data are included in this paper. Cohen's *d* effect sizes were calculated to supplement reported means and standard deviations of cognitive task scores before and after training within each training group. These data were reported for transparency and to support future meta-analyses rather than as a major effect of interest in this study.

2.6.2 Structural Analyses

Cortical area, thickness, and volume were compared before and after training and between groups across several regions of the frontal and parietal lobes, as well as the cingulate cortex and insular cortex. These anatomical areas were extracted from FreeSurfer image maps for region of interest (ROI) analysis and were based on the previously discussed structural and functional neuroimaging investigations and lesion studies demonstrating neural correlates of working memory. Frontal lobe ROIs were the superior, rostral middle, and caudal middle frontal gyri, the precentral gyrus, and three divisions of the inferior frontal gyrus: the pars opercularis, pars triangularis, and pars orbitalis. Parietal lobe regions included the postcentral and supramarginal gyri, and the superior and inferior parietal cortex. Regions of the cingulate cortex were the rostral and caudal anterior divisions, and within the insular cortex, the insula was examined. Volumes were also examined for several subcortical regions, specifically, the caudate, putamen, pallidum, thalamus, and hippocampus.

In addition to a priori ROIs, FreeSurfer software was used to calculate whole brain volumes for comparison between training groups. Total gray matter volume is a measure of all subcortical gray matter voxels plus cerebellum gray matter and the left and right cortex. White matter volume was calculated based on all white matter voxels below the cortical ribbon.

Resulting measurements were imported into SPSS. Structural neuroimaging data were analyzed using 2 (group) x 2 (time) RM-ANOVAs with alpha adjusted to <.001 given the large number of comparisons. Analyses were conducted separately for: 1) the effect of group and time on gray matter area in each left and right cortical ROI, 2) the effect of group and time on gray matter thickness in each left and right cortical ROI, and 3) the effect of group and time on gray matter volume in each left and right cortical and subcortical ROI as well as total gray and white matter volumes. Cohen's *d* effect sizes were calculated to supplement reported means and standard deviations of cortical area, thickness, and volume before and after training for each group.

Finally, correlations were planned for gray and white matter total volumes to identify associations between baseline factors and volume change after training, specifically (1) baseline volumes and volume change, (2) baseline age and volume change, and (3) baseline estimated intelligence and volume change.

Chapter Three: Results

3.1 Participant flow

Of the 359 individuals who expressed interest in participating in this study, a total of 187 were eligible and invited to participate. Of these, 107 individuals provided informed consent. From this group, a total of 66 participants were randomized to the working memory training group ($n = 33$) or the processing speed training group ($n = 33$). Remaining participants were assigned to a non-MRI group which was part of the larger study ($n = 29$) or either withdrew or were deemed ineligible prior to their first MRI appointment ($n = 12$). Three participants were removed from the study during or after scanning due to inability to tolerate the MRI scan ($n = 2$) or incidental findings found during the scanning session ($n = 1$). Screening, eligibility, consent, and completion rates for the working memory training and processing speed training groups are presented in Figure 1. Of the 63 individuals who completed baseline assessments and were randomized into a training group, 76% completed all components of the study (pre- and post-testing, successful MRI sessions, and completion of their assigned training program). Of the participants who failed to complete the study, primary reasons reported were loss of interest in the training program, and changes in internet access due to travel. Data from two participants from the working memory training group were removed from analysis due to poor training compliance and data for another working memory trainee were removed due to imaging acquisition issues.

There were no significant demographic, personality, motivational, mood, lifestyle, cognitive, or neuroanatomical (i.e., cortical area, thickness, or volume) differences at baseline between those who completed and those who did not complete the study.

3.2 Participant Characteristics

Participant mean age was 30.88 (*SD* = 5.97) years. They were mostly white (74%), well educated (M_{years} = 15.43, *SD* = 2.06), and worked full time (63%). Regarding gender and relationship status, 56% of participants identified as female and 43% reported being in a coupled relationship. Data on biological sex at birth were not collected.

Independent samples t-tests and chi-square tests revealed that there were no significant differences between the two groups at baseline on any demographic, personality, lifestyle, or cognitive variables. Regarding mood, participants assigned to the processing speed training group reported more fatigue at baseline (*M* = 1.50, *SD* = 0.92) relative to those assigned to working memory training (*M* = 0.98, *SD* = 0.80), $t_{46} = 2.12$, $p = 0.04$. No further mood differences were present at baseline. Importantly, there were no significant cortical area, thickness, or volume differences between the two groups at baseline in any cortical or subcortical regions of interest or in total gray or white matter volumes. Inspection of QQ plots on all variables revealed normal distributions at baseline. All baseline characteristics and comparisons are presented in Table 1.

Table 1.

Participant characteristics at baseline

	Working Memory Group Mean (SD)	Processing Speed Group Mean (SD)	Test(df)	*p*
<u>*Demographics*</u>				
N	24	24		
Age	30.43 (6.24)	31.33 (5.78)	$t_{46} = 0.52$	.61
Gender (% female)	54	58	χ^2 (1, *N* = 48) = 0.09	.77
Ethnicity (% White)	78	71	χ^2 (1, *N* = 48) = 0.34	.56
Marital status (% Coupled) ¶	38	50	χ^2 (1, *N* = 47) = 0.76	.38
Years of Education¶	15.29 (2.22)	15.57 (1.93)	$t_{45} = 0.45$	.65
Employment (% Full time)†	71	54	χ^2 (2, N = 48) = 1.43	.49
Income (% over $50,000/yr)±	71	71	χ^2 (2, N = 48) = 1.12	.57
<u>*Cognitive Ability*</u>				
WAIS-IV FSIQ	107.83 (16.14)	111.63 (12.35)	$t_{46} = 0.91$	.37
WAIS-IV Block Design	24.96 (5.65)	26.92 (5.65)	$t_{46} = 1.21$	.24

WAIS-IV Matrix Reasoning	10.79 (1.98)	11.04 (1.49)	$t_{46} = 0.50$	.62
RAPM	12.71 (3.02)	13.50 (2.27)	$t_{46} = 1.03$	.31
CCFT	26.75 (4.51)	28.46 (3.51)	$t_{46} = 1.46$	.15
WAIS-IV Digit Span	29.67 (4.91)	31.17 (5.26)	$t_{46} = 1.02$	.31
WAIS-IV Arithmetic	7.46 (2.41)	7.63 (2.10)	$t_{46} = 0.26$	.80
Aospan total[¶]	53.74 (13.99)	53.96 (12.69)	$t_{45} = 0.06$	.96
SDRT Maintenance	17.63 (1.28)	18.25 (1.39)	$t_{46} = 1.62$	.11
SDRT Manipulation	16.00 (2.25)	16.63 (2.10)	$t_{46} = 1.00$	.33
WAIS-IV Symbol Search	40.58 (8.40)	39.58 (5.79)	$t_{46} = 0.48$	.63
WAIS-IV Coding	81.29 (14.72)	85.25 (12.01)	$t_{46} = 1.02$	.31
Personality Factors				
Grit Score	3.49 (0.65)	3.31 (0.61)	$t_{46} = 1.00$	.33
Need for Cognition	69.71 (6.67)	69.06 (9.51)	$t_{46} = 0.28$	.79
HEXACO: Honesty-Humility	3.71 (0.48)	3.80 (0.54)	$t_{46} = 0.65$	.52
HEXACO: Emotionality	2.99 (0.68)	2.83 (0.80)	$t_{46} = 0.78$	.44
HEXACO: Extraversion	3.41 (0.51)	3.54 (0.76)	$t_{46} = 0.71$	.48
HEXACO: Agreeableness	3.20 (0.55)	3.37 (0.50)	$t_{46} = 1.13$	.26
HEXACO: Conscientiousness	3.83 (0.52)	3.53 (0.64)	$t_{46} = 1.78$	.08
HEXACO: Openness	3.63 (0.68)	3.93 (0.42)	$t_{46} = 1.77$	.08
Mood and Lifestyle Factors				
PSQI Sleep Quality (Good:Poor)	13:11	13:11	χ^2 (1, N = 48) = 0.00	1.00
IPAQ Total MET[¶]	588.89 (427.01)	581.39 (320.05)	$t_{42} = 0.07$	.95
POMS: Vigor	2.37 (0.79)	1.97 (0.79)	$t_{46} = 1.74$	.08
POMS: Confusion	0.84 (0.72)	0.85 (0.63)	$t_{46} = 0.04$	.97
POMS: Tension	1.19 (0.85)	0.96 (0.81)	$t_{46} = 0.98$	.34
POMS: Anger	0.59 (0.48)	0.53 (0.41)	$t_{46} = 0.45$	.65
POMS: Fatigue	0.98 (0.80)	1.50 (0.92)	$t_{46} = 2.12$	.04
POMS: Depression	0.57 (0.54)	0.35 (0.46)	$t_{46} = 1.47$	.15
Training Motivation and Expectations				
Pre-training motivation	5.67 (0.88)	5.77 (0.78)	$t_{46} = 0.43$	.67
Pre-training expected improvement	4.54 (0.98)	4.42 (1.10)	$t_{46} = 0.42$	.68

WAIS-IV = Wechsler Adult Intelligence Scale (4th Edition); FSIQ = Full Scale Intelligence Quotient; RAPM = Raven's Advanced Progressive Matrices; CCFT = Cattell's Culture Fair Test; Aospan = Automated Operation Span; SDRT = Spatial Delay Response Task; PSQI = Pittsburgh Sleep Quality Index; IPAQ = International Physical Activity Questionnaire; MET = multiple of metabolic rate indicating energy expenditure; POMS = Profile of Moods State questionnaire

[¶]Demographic data is missing from the working memory training group for ethnicity (n=1) and from the processing speed group for years of education (n=1). AOSPAN data for one working memory group participant's baseline score did not record. Cognitive activities data is missing for one working memory group participant and three processing speed group participants. Total IPAQ scores for two from each training group could not be calculated due to missing subscale scores.

[†]Employment data was tested using three categories: full-time, part-time, and unemployed.

[±] Income (household gross) data was tested using three categories: <$50,000/year, $50,000-$95,000/year, and >$95,000/year

3.3 Convergent Validity of Cognitive Tasks

Two-tailed correlations, presented in Table 2, demonstrate that the theoretically related tasks within each cognitive domain were statistically related in this sample. Tasks falling under the domain of perceptual reasoning were Block Design, Matrix Reasoning, CCFT, and RAPM and were moderately to strongly correlated: *r*s = .37 - 68, *p*s < .001 - < .01. Verbal working memory tasks were Aospan, Digit Span, and Arithmetic. Arithmetic was not significantly correlated with Aospan, $r = 0.14$, $p = .30$, perhaps reflecting differences in stimuli complexity (simple math versus more complex word problems) or delivery (auditory versus digital). However, Arithmetic and Digit Span were moderately correlated, $r = 32$, p < .05 as were Aospan and Digit Span, $r = .48$, p < .001. The two spatial working memory tasks were moderately correlated with each other, $r = .43$, $p < .001$ but were not strongly associated with the verbal working memory tasks. SDRT Manipulation was weakly, although significantly, associated with Digit Span, $r = .30$, $p < .05$, and Arithmetic, $r = .28$, $p < .05$, perhaps because those tasks require the examinee to maintain and work with the information held in mind, but not with Aospan, $r = .06$, $p = .65$ which relies on updating rather than manipulation. In contrast, SDRT Maintenance was not associated with any of the verbal working memory tasks, *r*s = -.03 - .14, *p*s = .19 - .83. The two processing speed tasks were moderately correlated, $r = .59$, $p < .001$.

Table 2.

Correlations among cognitive test scores at baseline

Measure:	Block Design	Matrix Reas.	CCFT	RAPM	Aospan	Digit Span	Arith.	SDRT Main.	SDRT Mani.	Symbol Search
Block Design	–									
Matrix Reas.	.43***	–								
CCFT	.56***	.37**	–							
RAPM	.68***	.48***	.62***	–						
Aospan	.32*	.23	.31*	.20	–					
Digit Span	.32*	.23	.44***	.21	.48***	–				
Arith.	.22	.24	.37**	.27*	.14	.32*	–			
SDRT Main.	.22	.19	.14	.21	-.03	.17	.14	–		
SDRT Mani.	.41**	.40**	.37**	.39**	.06	.30*	.28*	.43***	–	
Symbol Search	.32*	.05	.40**	.25*	.36**	.29*	.27*	-.15	.07	–
Coding	.24	.13	.36**	.13	.18	.39**	.15	.17	.08	.59***

CCFT = Cattell's Culture Fair Test; RAPM = Raven's Advanced Progressive Matrices; Aospan = Automated Operation Span; Arith. = Arithmetic; SDRT Main. = Spatial Delay Response Task Maintenance subtest; SDRT Mani. = Spatial Delay Response Task Manipulation subtest

* $p < .05$ ** $p < .01$ ** $p < .001$

3.4 Training characteristics

On average, working memory trainees spent 13.22 (*SD* = 4.76) hours actively training on assigned games, and processing speed trainees spent 11.69 (*SD* = 2.97) hours training. The two groups did not significantly differ in total training dosage. The working memory training group engaged in more non-assigned game-play (i.e., "contamination") with an average of 10.66 (*SD* = 14.76) minutes of contamination relative to 1.87 (*SD* = 6.43) minutes in the processing speed group, $t_{46} = 2.67$, $p = 0.01$. However, contamination in the working memory training group represented only 1.34% of the total training time. Furthermore, almost half (*M* = 5.00, *SD* = 6.84) of the contamination time recorded in the working memory training group was spent on a non-assigned task which targeted working memory (i.e., a visual memory span task).

Significant gains in individual training games were observed across the training period in both groups based on comparison of the average of the first five iterations of each game to the last five iterations of each game. Within the working memory training group, Memory Lane n-back level from the first five (M = 1.93, SD = 0.18) to the last five (M = 5.27, SD = 1.91) iterations significantly increased, (t_{23} = 8.54, p < .001). Similarly, correct matches in the first five (M = 27.69, SD = 6.20) to the last five (M = 54.69, SD = 6.75) iterations of Memory Match significantly increased, (t_{23} = 19.51, p < .001) and correct matches in the first five (M = 34.22, SD = 9.08) to the last five (M = 62.61, SD = 10.84) iterations of Memory Match Overload significantly increased, (t_{23} =11.31, p < .001). Gains across the working memory training period are depicted in Figure 4.

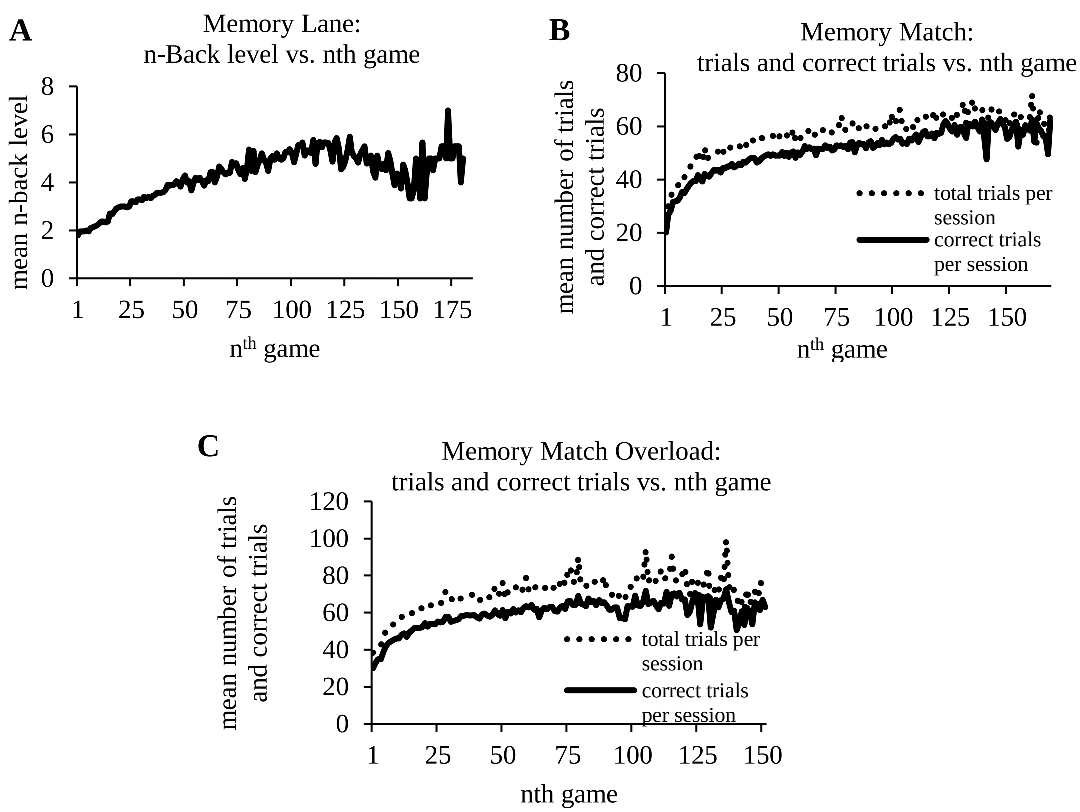

Figure 4. Mean performance by working memory training game. Higher scores indicates better performance.

Within the processing speed training group, Spatial Speed Match reaction times from the first five (*M* = 766.20, *SD* = 185.25) to the last five (*M* = 477.12, *SD* = 31.32) games significantly decreased, t_{23} = 8.03, *p* <.001. Speed Match reaction times from the first five (*M* = 689.30, *SD* = 110.07) to the last five (*M* = 472.68, *SD* = 26.21) games also significantly decreased, t_{23} = 9.74, *p* <.001. Finally, Speed Match Overdrive reaction times from the first five (*M* = 996.86, *SD* = 144.49) to the last five (*M* = 571.53, *SD* = 35.77) games significantly decreased, t_{23} = 14.31, *p* <.001. Reaction time decreases across the processing speed training period are shown in Figure 5.

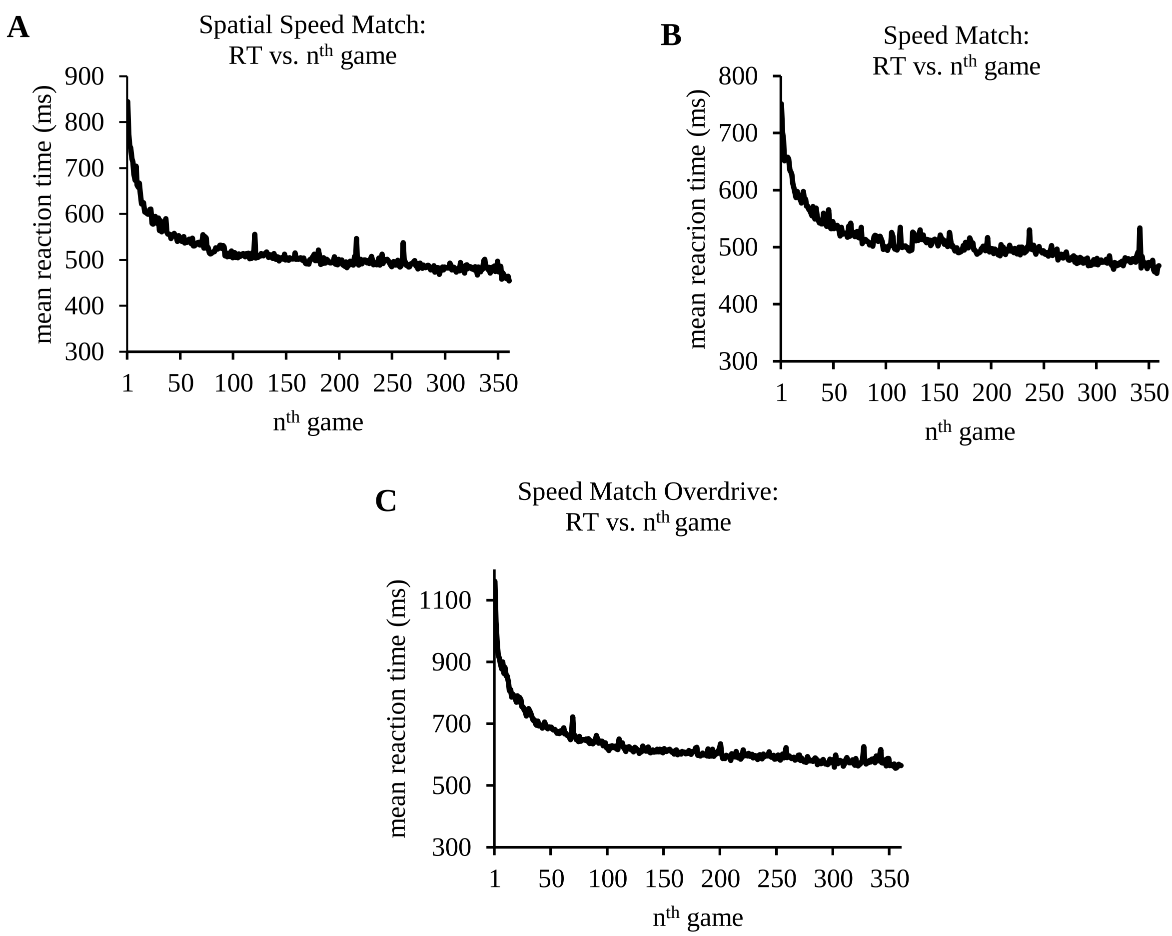

Figure 5. Mean performance by processing speed training game. Decreased reaction times indicate better performance.

3.5 Cognitive outcomes

The following results reflect that of the two training groups: the experimental group and the active control group. Although pre- and post-training data was collected from a no-contact control group as part of a larger study's protocol, those data are not presented here because no-contact controls did not undergo neuroimaging. Furthermore, data comparing cognitive task outcomes of the two training groups and the no-contact control group before and after the training period were recently reported in Clark and colleagues (2017c).

Cognitive outcomes for the training groups were analyzed separately for each task within each of the four cognitive domains, and alpha levels were adjusted based on the number of tests within each domain (p = 0.05/n). In the perceptual reasoning cognitive domain, the RM-ANOVA revealed a main effect of time for the Block Design task (alpha adjusted to 0.01), $F(1, 46) = 19.44$, $p < .001$. Paradoxically, scores were lower after training. No significant group or interaction effects emerged for the Block Design task. No main effects or interactions were present for remaining perceptual reasoning tasks.

Regarding the three verbal working memory tasks, the RM-ANOVA resulted in a main effect of time for Aospan (alpha adjusted to 0.016), $F(1, 45) = 9.69$, $p = .003$ with higher scores after training. No significant group or interaction effects emerged. No other main effects were present for verbal working memory tasks.

No main effects or interactions were revealed for either of the two spatial working memory tasks, using an adjusted alpha level of 0.025.

For the two processing speed tasks, a main effect of time was found for both Symbol Search and Coding (alpha adjusted to 0.025), $F(1, 46) = 7.86$, $p = .007$ and $F(1, 46) = 34.65$, $p < .001$, respectively. In both tasks there was strong increase in scores (denoting faster speed) after training. However, no significant effects of group and no interactions emerged for either task. Means and standard deviations of all cognitive outcomes are available in Table 3, and RM-ANOVA data are presented in Table 4.

Table 3.

Means, standard deviations, and effects of time on cognitive test scores before and after within working memory training (n = 24) and processing speed training (n = 24) groups

Task	Group	Mean (SD) Time 1	Mean (SD) Time 2	*d* [95% CI]
Perceptual reasoning				
Block Design	WM	24.96 (5.66)	22.96 (4.15)	-0.41 [-0.96, 1.79]
	PS	26.92 (5.62)	23.83 (4.27)	-0.63 [-0.75, 2.02]
Matrix Reasoning	WM	10.79 (1.98)	10.29 (1.64)	-0.28 [-0.22, 0.78]
	PS	11.04 (1.49)	10.58 (1.41)	-0.35 [-0.01, 0.72]
CCFT	WM	8.12 (1.19)	8.29 (2.05)	0.10 [-0.57, 0.36]
	PS	8.25 (1.26)	8.08 (1.23)	-0.14 [-0.21, 0.48]
RAPM	WM	12.71 (3.03)	11.96 (3.45)	-0.24 [-0.66, 1.14]
	PS	13.50 (2.27)	13.04 (2.29)	-0.21 [-0.43, 0.84]
Verbal working memory				
Aospan ¶	WM	53.74 (13.99)	57.26 (10.01)	0.30 [-3.69, 3.10]
	PS	53.96 (12.69)	60.58 (11.73)	0.30 [-3.46, 2.87]
Digit Span	WM	29.43 (4.89)	29.52 (4.93)	0.02 [-1.38, 1.34]
	PS	31.17 (5.26)	31.04 (3.82)	-0.03 [-1.24, 1.30]
Arithmetic	WM	7.35 (2.41)	7.09 (1.38)	-0.14 [-0.41, 0.68]
	PS	7.63 (2.10)	7.21 (2.04)	-0.21 [-0.37, 0.78]
Spatial working memory				
SDRT maintenance	WM	17.63 (1.28)	17.75 (1.70)	0.08 [-0.50, 0.34]
	PS	18.25 (1.39)	18.63 (1.38)	0.28 [-0.66, 0.10]
SDRT manipulation	WM	16.00 (2.25)	16.13 (1.85)	0.06 [-0.64, 0.51]
	PS	16.63 (2.10)	16.17 (2.18)	-0.22 [-0.37, 0.81]
Processing Speed				
Symbol Search	WM	40.58 (8.40)	42.58 (7.58)	0.26 [-2.47, 1.96]
	PS	39.58 (5.73)	41.58 (8.74)	0.28 [-2.32, 1.77]
Coding	WM	81.29 (14.72)	87.42 (14.57)	0.43 [-4.48, 3.63]
	PS	85.25 (12.01)	93.38 (11.94)	0.69 [-4.01, 2.62]

Aospan = Automated Operation Span; SM = spatial maintenance; SMM = spatial maintenance plus manipulation; WM = working memory training group, PS = processing speed training group, RAPM = Raven's Advanced Progressive Matrices, CCFT = Cattel's Culture Fair Fest

¶Task data for one working memory group participant's baseline score did not record therefore working memory group *n* = 23 for this task

Table 4.

Effect of group and time on cognitive test scores

Task	Group Effect F(df)	*p*	Time Effect F(df)	*p*	Interaction F(df)	*p*
Perceptual reasoning						
Block Design	1.16 (1,46)	.29	19.44(1,46)	<*.001*	0.88 (1,46)	.35
Matrix Reasoning	0.56 (1,46)	.46	2.39 (1,46)	.13	0.01 (1,46)	.95
CCFT	0.02 (1,46)	.90	0.00 (1,46)	1.00	0.33 (1,46)	.57
RAPM	1.63 (1,46)	.21	3.14 (1,46)	.08	0.18 (1,46)	.67
Verbal working memory						
Aospan¶	0.31 (1,45)	.58	9.69 (1,45)	.003	0.91 (1,45)	.35
Digit Span	1.77 (1,46)	.19	0.01 (1,46)	.98	0.03 (1,46)	.87
Arithmetic	0.17 (1,46)	.68	1.04 (1,46)	.31	0.06 (1,46)	.82
Spatial working memory						
SDRT Maintenance	5.21 (1,46)	.03	0.94 (1,46)	.34	0.24 (1,46)	.63
SDRT Manipulation	0.43 (1,46)	.52	0.26 (1,46)	.62	0.78 (1,46)	.38
Processing Speed						
Symbol Search	0.23 (1,46)	.64	7.86 (1,46)	.007	0.00 (1,46)	1.00
Coding	1.83 (1,46)	.18	34.65 (1,46)	<.001	0.68 (1,46)	.41

Aospan = Automated Operation Span; SM = spatial maintenance; SMM = spatial maintenance plus manipulation; WM = working memory training group, PS = processing speed training group, RAPM = Raven's Advanced Progressive Matrices, CCFT = Cattel's Culture Fair Fest
¶Task data for one working memory group participant's baseline score did not record therefore working memory group *n* = 23 for this task
* p < .01 ** p < .001

3.6 Structural outcomes

Means, standard deviations, and Cohen's *d* effect sizes of neuroanatomical measurements before (time 1) and after (time 2) training are presented in Table 5 for cortical (area, thickness, and volume measurements) and subcortical (volume measurements) regions of interest. Table 6 shows means, standard deviations, and Cohen's *d* effect sizes of total gray and white matter volumes before and after training. Using an alpha level of <.001 due to adjustment for multiple comparisons, RM-ANOVAs examining group (working memory training, processing speed training) and time (pre-, post) failed to reveal significant group, time, or interaction effects for cortical surface area (Table 7; *F*s = 0.00 - 5.85, *p*s = .02 - .99), thickness (Table 8; *F*s = 0.01 -

9.38, *p*s = .004 - .98), or volume changes in left or right frontal or parietal lobe regions of interest (Table 9; *F*s = 0.00 - 10.84; *p*s = .002 - .99), or in cingulate or insular cortical regions (Table 9; *F*s = 0.00 - 9.61, *p*s = .003 - .97). Similarly, no significant group, time, or interaction effects were revealed for volume estimates within subcortical regions of interest (Table 9; *F*s = 0.00 – 4.64, *p*s = .04 - .99). Accordingly, no main effects or interactions were identified for total subcortical gray matter volumes or total gray matter volumes (Table 9; *F*s = 0.00 – 4.35, *p*s = .04 - .98). Finally, no main effects or interactions emerged from analyses of total white matter volumes (Table 9; *F*s = 0.08 - 1.14, *p*s = .29 - .78).

When the data were considered without adjusting for multiple comparisons (alpha < .05), main effects of time were revealed for left supramarginal gyrus thickness, $F(1, 46) = 5.77$, $p = .02$, and volume estimates, $F(1, 46) = 5.92$, p = .02 driven by decreased thickness and volume after training. An interaction effect was revealed for left supramarginal gyrus volume $F(1, 46) = 5.65$, $p = .02$ driven by decreased volume in the processing speed training group (control condition), and unchanged volume in the working memory training group (experimental condition). Further interaction effects that emerged in cortical surface area (Table 7; *F*s = 4.78 - 5.21, *p*s = .03 - .02), thickness (Table 8; *F*s = 5.58 - 9.38 *p*s = .004 - .02), and volume (Table 9; *F*s = 4.35 - 10.84 *p*s = .002 - .04) were driven by slight increases (effect sizes close to zero) in measurements within the working memory training group combined with slight decreases (effect sizes close to zero) in the processing speed training (control) group.

Table 5.

Means and standard deviations of cortical area (mm²), thickness (mm), and volume (mm³) before and after working memory training (n = 24) and processing speed training (n = 24)

Region	Measure	Group	Left Hemisphere			Right Hemisphere		
			Mean (SD) Time 1	Mean (SD) Time 2	*d* [95% CI]	Mean (SD) Time 1	Mean (SD) Time 2	*d* [95% CI]
Frontal Lobe								
Superior frontal gyrus								
	Area	WM	7141.04 (732.96)	7142.83 (743.44)	0.002 [-204.45, 204.45]	6846.50 (601.25)	6850.54 (606.15)	0.007 [-167.19, 167.20]
		PS	7157.42 (717.49)	7118.92 (703.54)	-0.06 [-196.84, 196.72]	6982.79 (763.18)	6966.29 (754.95)	-0.02 [-210.20, 210.20]
	Thickness	WM	2.84 (0.14)	2.85 (0.14)	0.07 [0.03, 0.11]	2.82 (0.14)	2.83 (0.16)	0.07 [0.026, 0.11]
		PS	2.87 (0.13)	2.86 (0.14)	-0.07 [-0.11, -0.04]	2.87 (0.13)	2.84 (0.12)	-0.25 [-0.28, -0.21]
	Volume	WM	23784.50 (2637.32)	23833.46 (2769.06)	0.02 [-748.83, 748.86]	22604.54 (2124.39)	22727.83 (2280.74)	0.06 [-610.31, 610.42]
		PS	23795.79 (2724.92)	23691.83 (2706.67)	-0.04 [-752.46, 752.08]	23424.88 (2817.10)	23149.42 (2835.50)	-0.10 [-782.82, 782.62]
Rostral middle frontal gyrus								
	Area	WM	5684.88 (776.56)	5685.71 (774.73)	0.001 [-210.70, 210.70]	5761.71 (767.17)	5746.96 (786.86)	-0.02 [-215.23, 215.18]
		PS	5640.58 (800.80)	5638.00 (817.24)	-0.003 [-224.07, 224.06]	5705.79 (722.45)	5682.29 (752.73)	-0.03 [-204.34, 204.28]
	Thickness	WM	2.43 (0.10)	2.44 (0.12)	0.09 [0.06, 0.12]	2.43 (0.12)	2.44 (0.13)	0.08 [0.05, 0.12]
		PS	2.47 (0.13)	2.48 (0.14)	0.09 [0.06, 0.12]	2.46 (0.11)	2.45 (0.12)	-0.09 [-0.12, -0.06]
	Volume	WM	16646.92 (2488.88)	16676.67 (2599.99)	0.01 [-704.81, 704.84]	16758.50 (2474.29)	16781.08 (2658.43)	0.009 [-711.18, 711.19]
		PS	16700.92 (2834.81)	16783.54 (3060.19)	0.03 [-753.23, 753.29]	16808.17 (2186.30)	16701.58 (2327.90)	-0.05 [-625.44, 625.34]
Caudal middle frontal gyrus								

Area	WM	2371.13	2373.29	0.006	2168.33	2174.96	0.02
		(350.36)	(350.06)	[-96.98, 96.99]	(290.93)	(296.43)	[-81.31, 81.36]
	PS	2280.33	2272.50	-0.02	2148.92	2135.08	-0.04
		(424.43)	(407.62)	[-115.26, 115.22]	(365.51)	(349.29)	[-99.04, 98.96]
Thickness	WM	2.66	2.66	0.00	2.66	2.66	0.00
		(0.13)	(0.13)	[-0.22, 0.22]	(0.11)	(0.13)	[-0.22, 0.22]
	PS	2.67	2.66	-0.08	2.65	2.64	-0.10
		(0.13)	(0.12)	[-0.12, -0.05]	(0.10)	(0.11)	[-0.13, 0.07]
Volume	WM	7083.46	7090.46	0.006	6471.46	6529.00	0.06
		(1186.41)	(1220.95)	[-333.38, 333.39]	(945.31)	(981.31)	[-266.77, 266.89]
	PS	6832.50	6796.75	-0.03	6497.00	6436.63	-0.5
		(1362.93)	(1342.92)	[-374.72, 374.66]	(1197.62)	(1166.86)	[-327.49, 327.39]
Pars opercularis¶							
Area	WM	1643.33	1648.67	0.02	1393.08	1392.54	-0.002
		(232.36)	(235.33)	[-64.74, 64.79]	(238.72)	(238.89)	[-66.14, 66.13]
	PS	1694.13	1685.04	-0.04	1484.88	1475.88	-0.3
		(270.13)	(258.87)	[-73.30, 73.23]	(284.39)	(276.80)	[-77.75, 77.68]
Thickness	WM	2.66	2.66	0.00	2.72	2.72	0.00
		(0.13)	(0.13)	[-0.04, 0.04]	(0.15)	(0.17)	[-0.04, 0.04]
	PS	2.70	2.70	0.00	2.69	2.67	-0.20
		(0.11)	(0.10)	[-0.30, 0.30]	(0.10)	(0.10)	[-0.23, -0.18]
Volume	WM	5036.63	5057.38	0.03	4304.83	4302.58	-0.003
		(754.38)	(779.62)	[-212.41, 212.47]	(734.49)	(782.29)	[-210.14, 210.13]
	PS	5267.54	5249.17	-0.02	4548.00	4515.75	-0.03
		(825.49)	(810.88)	[-226.62, 226.57]	(957.54)	(973.38)	[-267.42, 267.35]
Pars triangularis¶							
Area	WM	1301.75	1303.00	0.006	1456.25	1450.46	-0.03
		(201.79)	(207.13)	[-56.62, 56.63]	(238.76)	(237.76)	[-66.01, 65.96]
	PS	1293.71	1295.58	0.01	1482.42	1479.42	-0.01
		(183.23)	(178.46)	[-52.74, 52.76]	(282.73)	(286.47)	[-78.83, 78.81]
Thickness	WM	2.52	2.54	0.16	2.57	2.57	0.00
		(0.13)	(0.13)	[0.12, 0.19]	(0.13)	(0.13)	[-0.04, 0.04]
		2.57	2.57	0.00	2.59	2.58	-0.09
	PS	(0.12)	(0.13)	[-0.04, 0.04]	(0.13)	(0.11)	[-0.12, -0.05]
Volume	WM	3843.79	3874.79	0.05	4351.92	4351.54	0.00
		(641.84)	(685.71)	[-183.88, 183.97]	(731.63)	(745.90)	[-204.60, 204.60]
	PS	3901.75	3906.75	0.01	4481.63	4471.13	-0.01
		(641.67)	(624.34)	[-175.31, 175.33]	(926.46)	(929.79)	[-257.05, 257.02]

Pars orbitalis¶							
Area	WM	612.46	608.63	-0.04	778.83	775.42	0.03
		(90.76)	(87.54)	[-27.74, 24.65]	(125.59)	(124.35)	[-34.64, 34.58]
	PS	624.46	624.50	0.00	757.63	751.04	-0.07
		(84.64)	(87.29)	[-23.81, 23.81]	(104.64)	(103.48)	[-28.88, 28.75]
Thickness	WM	2.71	2.72	-0.05	2.80	2.80	0.00
		(0.23)	(0.21)	[-0.11, 0.02]	(0.18)	(0.16)	[-0.05, 0.05]
	PS	2.80	2.80	0.00	2.80	2.79	-0.05
		(0.16)	(0.17)	[-0.05, 0.05]	(0.20)	(0.20)	[-0.11, -0.001]
Volume	WM	2208.38	2211.75	0.10	2782.92	2784.25	0.003
		(331.25)	(329.35)	[-91.46, 91.48]	(475.11)	(451.12)	[-128.30, 128.30]
	PS	2332.50	2325.50	-0.02	2727.67	2704.33	-0.06
		(347.70)	(354.24)	[-97.22, 97.18]	(389.35)	(392.62)	[-108.34, 108.22]
Precentral gyrus							
Area	WM	4755.83	4750.79	-0.01	4899.13	4898.71	-0.00
		(461.02)	(457.71)	[-127.23, 127.21]	(437.27)	(435.97)	[-120.92, 120.92]
	PS	4717.63	4712.96	-0.01	4806.33	4800.63	-0.01
		(573.52)	(562.13)	[-157.27, 157.25]	(513.37)	(515.39)	[-142.46, 142.44]
Thickness	WM	2.67	2.68	0.11	2.65	2.64	-0.11
		(0.08)	(0.10)	[0.09, 0.14]	(0.09)	(0.09)	[-0.14, -0.09]
	PS	2.65	2.64	-0.09	2.62	2.61	-0.10
		(0.11)	(0.12)	[-0.12, -0.06]	(0.11)	(0.10)	[-0.13, -0.07]
Volume	WM	14136.04	14154.46	0.01	14394.21	14388.75	-0.00
		(1582.56)	(1678.66)	[-451.77, 451.79]	(1509.04)	(1496.15)	[-416.14, 416.13]
	PS	13798.17	13767.83	-0.02	13799.88	13747.08	-0.03
		(1742.55)	(1806.12)	[-497.48, 491.45]	(1625.31)	(1606.82)	[-447.59, 447.53]
Rostral anterior cingulate cortex							
Area	WM	849.21	844.71	-0.04	719.67	719.21	-0.01
		(112.49)	(117.04)	[-31.83, 31.75]	(97.45)	(98.60)	[-27.15, 27.14]
	PS	831.38	831.63	-0.002	695.00	682.33	-0.10
		(158.36)	(159.03)	[-43.95, 43.95]	(127.58)	(130.05)	[-35.78, 35.58]
Thickness	WM	2.99	3.00	0.06	3.08	3.10	0.08
		(0.16)	(0.18)	[-0.01, 0.12]	(0.25)	(0.26)	[0.01, 0.15]
	PS	3.08	3.10	0.08	3.10	3.10	0.00
		(0.24)	(0.27)	[0.01, 0.15]	(0.23)	(0.22)	[-0.06, 0.06]
Volume	WM	2842.46	2827.33	-0.04	2466.75	2502.04	0.10
		(379.36)	(408.22)	[-109.17, 109.10]	(338.73)	(351.49)	[-95.49, 95.70]

	PS	2884.92	2917.54	-0.06	2397.42	2354.25	-0.09
		(529.64)	(586.24)	[-154.66, 154.77]	(485.27)	(479.00)	[-133.62, 133.43]
<u>Caudal anterior cingulate cortex</u>							
Area	WM	657.08	656.13	-0.01	792.21	794.83	0.02
		(126.19)	(127.37)	[-35.12, 35.10]	(131.11)	(132.38)	[-36.47, 36.51]
	PS	668.17	664.25	-0.02	764.17	761.67	-0.01
		(170.12)	(173.82)	[-47.65, 47.61]	(178.90)	(173.54)	[-48.82, 48.79]
Thickness	WM	2.79	2.80	-0.04	2.71	2.73	0.06
		(0.28)	(0.27)	[-0.04, 0.11]	(0.32)	(0.35)	[-0.03, 0.15]
	PS	2.74	2.77	0.10	2.71	2.69	-0.06
		(0.32)	(0.32)	[-0.01, 0.18]	(0.33)	(0.31)	[-0.15, 0.03]
Volume	WM	1970.79	1961.46	-0.02	2350.88	2382.67	0.08
		(417.45)	(416.54)	[-115.51, 115.46]	(403.60)	(413.66)	[-133.10, 113.25]
	PS	1984.67	1995.13	-0.02	2328.92	2274.04	-0.09
		(578.10)	(597.86)	[-162.84, 162.88]	(644.98)	(593.43)	[-171.72, 171.54]
<u>*Parietal Lobe*</u>							
<u>Postcentral gyrus</u>							
Area	WM	4047.79	4036.04	-0.04	3880.08	3868.00	-0.04
		(320.87)	(314.72)	[-88.05, 87.98]	(323.40)	(313.69)	[-88.27, 88.20]
	PS	3978.38	3984.33	0.01	3810.42	3798.42	-0.03
		(519.11)	(511.76)	[-142.74, 142.76]	(489.86)	(482.47)	[-134.67, 134.62]
Thickness	WM	2.20	2.21	0.17	2.18	2.20	0.26
		(0.06)	(0.06)	[0.15, 0.19]	(0.08)	(0.08)	[0.23, 0.28]
	PS	2.23	2.21	-0.19	2.19	2.17	-0.19
		(0.11)	(0.11)	[-0.22, -0.16]	(0.11)	(0.11)	[-0.22, -0.16]
Volume	WM	10219.21	10241.25	0.03	9625.83	9686.29	0.06
		(896.59)	(939.78)	[-254.33, 254.38]	(934.56)	(987.82)	[-266.23, 266.36]
	PS	10199.29	10106.63	-0.06	9483.92	9381.75	-0.09
		(1661.29)	(1604.97)	[-452.40, 452.29]	(1233.82)	(1213.77)	[-339.02, 338.84]
<u>Supramarginal gyrus</u>							
Area	WM	3741.79	3745.38	0.01	3542.83	3540.63	-0.01
		(488.73)	(483.84)	[-134.67, 134.68]	(481.62)	(478.02)	[-132.89, 132.88]
	PS	3974.50	3951.54	-0.04	3461.21	3443.46	-0.03
		(625.32)	(605.31)	[-170.47, 170.39]	(638.08)	(621.24)	[-174.42, 174.37]
Thickness	WM	2.67	2.66	-0.10	2.66	2.66	0.00
		(0.10)	(0.11)	[-0.13, -0.07]	(0.08)	(0.10)	[-0.03, 0.03]
	PS	2.68	2.66	-0.17	2.69	2.66	-0.22
		(0.12)	(0.12)	[-0.20, -0.14]	(0.15)	(0.13)	[-0.26, -0.18]

Volume	WM	11150.75	11149.42	-0.00	10411.21	10424.08	0.01
		(1408.87)	(1403.88)	[-389.48, 389.48]	(1520.44)	(1551.94)	[-425.45, 425.46]
	PS	11878.13	11764.83	-0.07	10246.46	10147.96	-0.05
		(1819.66)	(1744.91)	[-493.76, 493.63]	(1873.91)	(1839.86)	[-514.32, 514.21]
Superior parietal cortex							
Area	WM	5208.79	5214.33	-0.14	5240.00	5235.58	-0.10
		(413.12)	(420.39)	[-115.41, 115.43]	(444.27)	(446.16)	[-123.31, 123.29]
	PS	5255.00	5231.04	-0.04	5253.29	5227.00	-0.05
		(591.58)	(580.01)	[-162.28, 162.20]	(513.31)	(494.94)	[-139.69, 139.58]
Thickness	WM	2.27	2.27	0.00	2.27	2.28	0.12
		(0.08)	(0.08)	[-0.2, 0.2]	(0.08)	(0.09)	[0.10, 0.14]
	PS	2.30	2.28	-0.24	2.30	2.27	-0.29
		(0.08)	(0.09)	[-0.26, -0.22]	(0.10)	(0.11)	[-0.32, -0.26]
Volume	WM	13463.50	13495.96	0.03	13551.96	13596.88	0.03
		(1130.53)	(1204.16)	[-323.42, 323.47]	(1420.66)	(1433.04)	[-395.12, 395.19]
	PS	13595.92	13435.21	-0.10	13723.33	13489.58	-0.17
		(1659.63)	(1640.98)	[-457.14, 456.94]	(1457.50)	(1386.25)	[-394.07, 393.73]
Inferior parietal cortex							
Area	WM	4524.13	4519.5	-0.01	5427.46	5424.75	-0.00.
		(647.91)	(651.21)	[-179.90, 179.88]	(869.34)	(871.86)	[-241.11, 241.10]
	PS	4427.38	4402.33	-0.04	5560.71	5539.33	-0.03
		(662.35)	(647.33)	[-181.40, 181.32]	(698.97)	(689.12)	[-192.25, 192.18]
Thickness	WM	2.50	2.51	-0.10	2.54	2.54	0.00
		(0.10)	(0.11)	[0.07, 0.13]	(0.11)	(0.13)	[-0.03, 0.03]
	PS	2.56	2.55	-0.09	2.57	2.55	-0.19
		(0.12)	(0.11)	[-0.12, -0.06]	(0.11)	(0.11)	[-0.22, -0.16]
Volume	WM	12829.25	12858.79	0.01	15839.13	15823.71	-0.01
		(2130.31)	(2259.91)	[-608.17, 608.19]	(2598.03)	(2662.45)	[-728.48, 728.47]
	PS	12835.29	12746.13	-0.05	16370.33	16209.29	-0.06
		(1910.88)	(1932.32)	[-532.22, 532.13]	(2636.83)	(2577.93)	[-722.20, 722.07]
Insular cortex							
Insula							
Area	WM	2095.63	2101.17	0.03	2190.00	2197.58	0.03
		(178.83)	(188.12)	[-50.80, 50.86]	(261.57)	(256.81)	[-71.75, 71.81]
	PS	2131.21	2122.13	-0.04	2210.63	2186.25	-0.07
		(263.49)	(266.65)	[-73.45, 73.38]	(363.07)	(353.95)	[-99.36, 99.22]
Thickness	WM	3.15	3.16	0.07	3.10	3.10	0.00
		(0.15)	(0.16)	[0.02, 0.11]	(0.17)	(0.17)	[-0.05, 0.05]

		PS	3.16	3.17	0.07	3.12	3.12	0.00
			(0.15)	(0.15)	[0.03, 0.11]	(0.16)	(0.15)	[-0.04, 0.04]
	Volume	WM	6693.38	6741.25	0.07	6868.29	6898.42	0.04
			(660.69)	(684.12)	[-186.17, 186.32]	(719.44)	(761.97)	[-205.17, 205.26]
		PS	6824.04	6802.63	-0.02	6989.25	6907.17	-0.6
			(1085.57)	(1081.07)	[-300.04, 300.00]	(1358.75)	(1319.66)	[-370.98, 370.86]
Subcortical Regions								
Caudate								
	Volume	WM	3951.39	3937.46	-0.03	4007.60	4020.26	0.02
			(560.55)	(563.43)	[-155.66, 155.61]	(557.41)	(559.33)	[-154.61, 154.66]
		PS	3887.07	3855.63	-0.05	3836.48	3805.32	-0.05
			(680.10)	(653.46)	[-187.74, 184.65]	(639.02)	(620.56)	[-174.48, 174.38]
Putamen								
	Volume	WM	5843.49	5850.75	0.01	5679.32	5654.63	-0.04
			(811.91)	(777.30)	[-220.10, 220.12]	(688.82)	(631.13)	[-182.99, 182.91]
		PS	6008.88	5958.38	-0.07	5732.35	5717.88	-0.02
			(714.99)	(677.39)	[-192.95, 192.80]	(728.76)	(631.13)	[-188.81, 188.77]
Pallidum								
	Volume	WM	1436.55	1427.12	-0.04	1586.77	1608.87	0.10
			(255.40)	(265.44)	[-72.17, 72.10]	(223.76)	(244.77)	[-64.85, 65.04]
		PS	1393.57	1406.93	0.07	1594.29	1601.44	0.04
			(217.60)	(203.29)	[-58.25, 58.38]	(201.76)	(198.27)	[-55.36, 55.43]
Thalamus								
	Volume	WM	8444.29	8445.74	0.002	7299.85	7332.35	0.05
			(925.48)	(923.42)	[-256.02, 256.02]	(723.06)	(744.51)	[-203.19, 203.28]
		PS	8671.15	8642.83	-0.03	7510.55	7468.93	-0.05
			(1014.49)	(1054.10)	[-286.52, 286.46]	(873.79)	(893.31)	[-244.75, 244.66]
Hippocampus								
	Volume	WM	4328.77	4332.93	0.01	4442.96	4438.88	-0.01
			(385.43)	(366.17)	[-104.10, 104.12]	(345.14)	(341.11)	[-95.04, 95.02]
		PS	4372.80	4370.90	-0.004	4534.82	4541.84	0.01
			(470.67)	(481.51)	[-131.86, 131.85]	(541.17)	(547.16)	[-150.69, 150.72]

¶Pars opercularis, pars triangularis, and pars orbitalis are divisions of the inferior frontal gyrus.

Table 6.

Means and standard deviations of total gray and white matter volumes (mm^3) before and after working memory training (n = 24) and processing speed training (n = 24)

Region	Group	Mean (SD) Time 1	Mean (SD) Time 2	*d* [95% CI]
Subcortical Gray Matter	WM	61601.21 (5363.12)	61588.00 (5322.86)	-0.003 [-1479.70, 1479.70]
	PS	62341.17 (5716.40)	62139.42 (5675.86)	-0.04 [-1577.53, 1577.46]
Total Gray Matter	WM	659196.20 (59263.23)	660046.54 (61983.47)	0.01 [-16793.24, 16793.27]
	PS	661791.74 (67830.03)	658197.67 (67881.02)	-0.05 [-18791.97, 18791.86]
Total White Matter	WM	453487.23 (54407.79)	453425.26 (54625.41)	-0.00 [-15097.86, 15097.86]
	PS	448569.78 (50477.45)	449617.42 (50347.24)	0.02 [-13961.19, 13961.23]

Table 7.

Effect of group and time on gray matter area (mm^2) for left and right hemisphere cortical regions of interest

Region	Group Effect F(df)	*p*	Time Effect F(df)	*p*	Interaction F(df)	*p*
<u>*Frontal Lobe*</u>						
L superior frontal gyrus	0.00 (1,46)	.99	2.28 (1,46)	.14	2.75 (1,46)	.10
R superior frontal gyrus	0.41 (1,46)	.53	0.34 (1,46)	.56	0.92 (1,46)	.34
L rostral middle frontal gyrus	0.04 (1,46)	.84	0.01 (1,46)	.94	0.02 (1,46)	.89
R rostral middle frontal gyrus	0.08 (1,46)	.78	2.52 (1,46)	.12	0.13 (1,46)	.72
L caudal middle frontal gyrus	0.75 (1,46)	.39	0.42 (1,46)	.52	1.32 (1,46)	.26
R caudal middle frontal gyrus	0.10 (1,46)	.76	0.43 (1,46)	.52	3.45 (1,46)	.07
L pars opercularis	0.37 (1,46)	.55	0.32 (1,46)	.57	4.78 (1,46)	.03
R pars opercularis	1.36 (1,46)	.25	2.17 (1,46)	.15	1.70 (1,46)	.20
L pars triangularis	0.02 (1,46)	.89	0.22 (1,46)	.60	0.01 (1,46)	.92
R pars triangularis	0.13 (1,46)	.72	2.24 (1,46)	.14	0.23 (1,46)	.64
L pars orbitalis	0.31 (1,46)	.58	0.44 (1,46)	.51	0.46 (1,46)	.50
R pars orbitalis	0.47 (1,46)	.50	5.61 (1,46)	.02	0.56 (1,46)	.46
L precentral gyrus	0.07 (1,46)	.80	0.51 (1,46)	.48	0.00 (1,46)	.98
R precentral gyrus	0.48 (1,46)	.49	0.19 (1,46)	.67	0.14 (1,46)	.71
<u>*Parietal Lobe*</u>						
L postcentral gyrus	0.24 (1,46)	.63	0.28 (1,46)	.60	2.60 (1,46)	.11
R postcentral gyrus	0.35 (1,46)	.56	3.38 (1,46)	.07	0.00 (1,46)	.99
L supramarginal gyrus	1.88 (1,46)	.18	2.78 (1,46)	.10	5.21 (1,46)	.03
R supramarginal gyrus	0.31 (1,46)	.58	3.62 (1,46)	.06	2.20 (1,46)	.15
L superior parietal cortex	0.05 (1,46)	.83	1.41 (1,46)	.24	3.61 (1,46)	.06
R superior parietal cortex	0.00 (1,46)	.99	4.26 (1,46)	.05	2.16 (1,46)	.15

L inferior parietal cortex	0.32 (1,46)	.57	5.85 (1,46)	.02	2.79 (1,46)	.10
R inferior parietal cortex	0.30 (1,46)	.59	2.11 (1,46)	.15	1.27 (1,46)	.27
Cingulate cortex						
L rostral anterior division	0.15 (1,46)	.70	0.33 (1,46)	.57	0.41 (1,46)	.53
R rostral anterior division	0.88 (1,46)	.35	3.41 (1,46)	.07	2.95 (1,46)	.09
L caudal anterior division	0.05 (1,46)	.83	0.51 (1,46)	.48	0.19 (1,46)	.67
R caudal anterior division	0.47 (1,46)	.50	0.00 (1,46)	.98	1.13 (1,46)	.29
Insular cortex						
L insula	0.19 (1,46)	.67	0.15 (1,46)	.70	2.59 (1,46)	.12
R insula	0.00 (1,46)	.96	1.32 (1,46)	.26	4.79 (1,46)	.03

**p* < .001

Table 8.

Effect of group and time on gray matter thickness (mm) for left and right hemisphere cortical regions of interest

Region	Group Effect F(df)	*p*	Time Effect F(df)	*p*	Interaction F(df)	*p*
Frontal Lobe						
L superior frontal gyrus	0.30 (1,46)	.58	0.01 (1,46)	.95	1.35 (1,46)	.25
R superior frontal gyrus	0.43 (1,46)	.54	1.28 (1,46)	.26	5.58 (1,46)	.02
L rostral middle frontal gyrus	1.43 (1,46)	.24	0.56 (1,46)	.46	0.02 (1,46)	.88
R rostral middle frontal gyrus	0.25 (1,46)	.62	0.09 (1,46)	.76	0.26 (1,46)	.62
L caudal middle frontal gyrus	0.09 (1,46)	.77	1.16 (1,46)	.29	0.14 (1,46)	.71
R caudal middle frontal gyrus	0.27 (1,46)	.61	0.15 (1,46)	.70	1.44 (1,46)	.24
L pars opercularis	1.43 (1,46)	.24	0.03 (1,46)	.86	0.14 (1,46)	.71
R pars opercularis	1.09 (1,46)	.30	1.65 (1,46)	.21	0.74 (1,46)	.39
L pars triangularis	1.39 (1,46)	.25	1.02 (1,46)	.32	0.49 (1,46)	.49
R pars triangularis	0.09 (1,46)	.77	0.69 (1,46)	.41	0.05 (1,46)	.83
L pars orbitalis	2.59 (1,46)	.11	0.24 (1,46)	.62	0.06 (1,46)	.81
R pars orbitalis	0.01 (1,46)	.94	0.08 (1,46)	.78	0.23 (1,46)	.63
L precentral gyrus	0.77 (1,46)	.39	0.29 (1,46)	.60	0.83 (1,46)	.37
R precentral gyrus	1.45 (1,46)	.24	1.09 (1,46)	.30	0.56 (1,46)	.46
Parietal Lobe						
L postcentral gyrus	0.14 (1,46)	.71	1.57 (1,46)	.22	7.38 (1,46)	.01
R postcentral gyrus	0.02 (1,46)	.88	0.11 (1,46)	.75	9.38 (1,46)	.004
L supramarginal gyrus	0.05 (1,46)	.83	5.77 (1,46)	.02	0.59 (1,46)	.45
R supramarginal gyrus	0.40 (1,46)	.53	4.18 (1,46)	.05	3.31 (1,46)	.08
L superior parietal cortex	0.83 (1,46)	.37	1.19 (1,46)	.28	2.36 (1,46)	.13
R superior parietal cortex	0.24 (1,46)	.63	2.15 (1,46)	.15	7.00 (1,46)	.01
L inferior parietal cortex	2.90 (1,46)	.10	0.45 (1,46)	.51	1.25 (1,46)	.27
R inferior parietal cortex	0.50 (1,46)	.48	3.22 (1,46)	.08	2.30 (1,46)	.14
Cingulate cortex						
L rostral anterior division	2.14 (1,46)	.15	0.84 (1,46)	.36	0.13 (1,46)	.72
R rostral anterior division	0.04 (1,46)	.83	0.66 (1,46)	.42	0.77 (1,46)	.39
L caudal anterior division	0.17 (1,46)	.68	1.08 (1,46)	.30	0.58 (1,46)	.45
R caudal anterior division	0.05 (1,46)	.83	0.06 (1,46)	.81	4.02 (1,46)	.05
Insular cortex						
L insula	0.00 (1,46)	.98	0.59 (1,46)	.45	0.89 (1,46)	.35
R insula	0.15 (1,46)	.70	0.01 (1,46)	.92	0.17 (1,46)	.68

**p* < .001

Table 9.

Effect of group and time on gray matter volume (mm^3) for left and right hemisphere cortical and subcortical regions of interest and total gray and white matter volumes

Region	Group Effect F(df)	*p*	Time Effect F(df)	*p*	Interaction F(df)	*p*
Frontal Lobe						
L superior frontal gyrus	0.01 (1,46)	.93	0.16 (1,46)	.69	1.27 (1,46)	.27
R superior frontal gyrus	0.73 (1,46)	.40	0.92 (1,46)	.34	6.32 (1,46)	.02
L rostral middle frontal gyrus	0.01 (1,46)	.92	0.68 (1,46)	.41	0.15 (1,46)	.70
R rostral middle frontal gyrus	0.00 (1,46)	.98	0.47 (1,46)	.50	1.10 (1,46)	.30
L caudal middle frontal gyrus	0.54 (1,46)	.47	0.81 (1,46)	.37	1.79 (1,46)	.19
R caudal middle frontal gyrus	0.01 (1,46)	.92	0.01 (1,46)	.95	6.22 (1,46)	.02
L pars opercularis	0.86 (1,46)	.36	0.01 (1,46)	.93	2.21 (1,46)	.14
R pars opercularis	0.83 (1,46)	.37	1.58 (1,46)	.21	1.20 (1,46)	.28
L pars triangularis	0.06 (1,46)	.81	2.35 (1,46)	.13	1.23 (1,46)	.27
R pars triangularis	0.27 (1,46)	.61	0.12 (1,46)	.73	0.11 (1,46)	.74
L pars orbitalis	1.47 (1,46)	.23	0.06 (1,46)	.81	0.47 (1,46)	.50
R pars orbitalis	0.30 (1,46)	.59	0.83 (1,46)	.37	1.04 (1,46)	.31
L precentral gyrus	0.54 (1,46)	.46	0.04 (1,46)	.85	0.62 (1,46)	.44
R precentral gyrus	1.89 (1,46)	.18	0.89 (1,46)	.35	0.59 (1,46)	.45
Parietal Lobe						
L postcentral gyrus	0.04 (1,46)	.84	2.24 (1,46)	.14	5.90 (1,46)	.02
R postcentral gyrus	0.50 (1,46)	.49	0.71 (1,46)	.40	10.84 (1,46)	.002
L supramarginal gyrus	2.10 (1,46)	.15	5.92 (1,46)	.02	5.65 (1,46)	.02
R supramarginal gyrus	0.20 (1,46)	.66	3.35 (1,46)	.07	5.67 (1,46)	.02
L superior parietal cortex	0.01 (1,46)	.93	1.53 (1,46)	.22	3.46 (1,46)	.07
R superior parietal cortex	0.01 (1,46)	.94	2.88 (1,46)	.10	6.26 (1,46)	.02
L inferior parietal cortex	0.01 (1,46)	.93	0.62 (1,46)	.44	2.45 (1,46)	.12
R inferior parietal cortex	0.37 (1,46)	.55	3.64 (1,46)	.06	2.48 (1,46)	.12
Cingulate cortex						
L rostral anterior division	0.23 (1,46)	.63	0.25 (1,46)	.62	1.89 (1,46)	.18
R rostral anterior division	0.82 (1,46)	.37	0.08 (1,46)	.78	7.69 (1,46)	.008
L caudal anterior division	0.03 (1,46)	.87	0.00 (1,46)	.97	0.44 (1,46)	.51
R caudal anterior division	0.19 (1,46)	.67	0.68 (1,46)	.41	9.61 (1,46)	.003
Insular cortex						
L insula	0.14 (1,46)	.71	0.53 (1,46)	.47	3.63 (1,46)	.06
R insula	0.04 (1,46)	.84	1.75 (1,46)	.19	8.16 (1,46)	.006
Subcortical regions						
L caudate	0.17 (1,46)	.68	4.26 (1,46)	.05	0.63 (1,46)	.43
R caudate	1.27 (1,46)	.27	0.60 (1,46)	.44	3.37 (1,46)	.07
L putamen	0.41 (1,46)	.53	0.90 (1,46)	.35	1.60 (1,46)	.21
R putamen	0.09 (1,46)	.77	1.25 (1,46)	.27	0.09 (1,46)	.77
L pallidum	0.22 (1,46)	.64	0.03 (1,46)	.86	1.10 (1,46)	.30
R pallidum	0.00 (1,46)	.99	1.73 (1,46)	.20	0.45 (1,46)	.51
L thalamus	0.56 (1,46)	.46	0.53 (1,46)	.47	0.65 (1,46)	.43
R thalamus	0.55 (1,46)	.46	0.07 (1,46)	.79	4.64 (1,46)	.04
L hippocampus	0.11 (1,46)	.74	0.01 (1,46)	.93	0.06 (1,46)	.82
R hippocampus	0.55 (1,46)	.46	0.02 (1,46)	.88	0.32 (1,46)	.58
Total volumes						
Subcortical gray matter	0.16 (1,46)	.69	1.91 (1,46)	.17	1.47 (1,46)	.23
Total gray matter	0.00 (1,46)	.98	1.66 (1,46)	.21	4.35 (1,46)	.04
Total white matter	0.08 (1,46)	.78	0.90 (1,46)	.35	1.14 (1,46)	.29

$^*p < .001$

3.7 Associations among Baseline Factors and Structural Outcomes

Regarding our secondary hypotheses, two-tailed correlations revealed that baseline age, estimated intelligence (i.e., IQ), and overall gray and white matter volumes were not associated with gray or white matter volume changes after working memory training: *r*s = -.32-.35, *p*s = .11-.83.

Chapter Four: Discussion

The field of working memory training is hindered by methodological inconsistency resulting in polarized findings and ambiguous conclusions. With methodological rigour based on recommendations stemming from the field, this study was designed in an attempt to replicate previous cognitive findings and extend the search for training-related benefits into the domain of brain structure as measured by MRI. Specifically, as part of a group of studies, we randomized healthy community adults, ranging in age from 18-40 years, into a six-week n-back working memory training (experimental condition) program or a six-week processing speed (active control condition) training program. A no-contact control group was included for analysis of cognitive tasks, but is not a part of this dissertation and was published by Clark and colleagues (2017c). In this dissertation, I compared behavioural and structural outcomes between the two groups over the training period, with hypotheses emphasizing potential changes in brain structure associated with training. This is one of only a few studies to examine gray and white matter structural plasticity associated with working memory training in healthy young adult brains. Furthermore, this is the first to include metrics beyond gray matter volume, namely, thickness and surface area, in a single investigation. Overall findings were consistent with the subset of literature suggesting that working memory training does not result in far transfer to cognitive outcomes, and null results were present across all structural metrics after correction for multiple comparisons.

4.1 Gray Matter Changes Associated with Working Memory Training

The present results are generally consistent with that of two comparable prior investigations of gray matter volume changes associated with working memory training. Similar to Metzler-Baddeley and colleagues (2016), I observed larger right caudal middle frontal gyrus volumes in the working memory training group, relative to the processing speed group, after training; however, as with Metzler-Baddeley and colleagues (2016), findings were not robust to correction for multiple comparison. Of note is that Metzler-Baddeley and colleagues' (2016) training condition was primarily span based, a condition that allows the use of strategy and therefore may have prevented adequate engagement of domain general working memory processes (Morrison & Chein, 2011). Colom and colleagues (2016) used a complex working memory task, specifically, the n-back training task similar to that used in the present study, and revealed significantly increased gray matter volume in the left cingulate cortex of trainees. In my study, larger volumes were revealed after working memory training in the right cingulate cortex, but again, these findings were not maintained when corrected for multiple comparisons. Of note is that the comparison group in Colom and colleagues' (2016) study was passive, rather than active; therefore, larger effect sizes are to be expected relative to our study that compared an experimental group to an active control group (Melby-Lervåg & Hulme, 2016). Furthermore, Colom and colleagues (2016) investigated larger structural areas, for example, entire lobes or structures, rather than specific regions of interest within those areas. Larger regions mean fewer tests; therefore, significant findings were not subjected to strict alpha adjustment due to multiple comparisons. Rather, $p<.05$ was retained as the standard used to deem a detected difference as statistically significant (Colom et al., 2016).

Regarding cortical thickness, one prior study examined changes in cortical thickness associated with working memory training. Román and colleagues (2016) described greater thickness in right ventral frontal areas of working memory trainees relative to passive controls;

however, they explained the difference as preserved gray matter thickness in trainees relative to observed thinning in passive controls. In my study, time by group interaction effects emerged in select frontal and parietal areas; however, statistical significance was not robust to correction for multiple comparisons. Furthermore, examination of mean scores revealed, consistent with Román and colleagues (2016), small increases in working memory trainees' cortical thickness measurements and small decreases in that of control trainees. Román and colleagues (2016) interpret such a seemingly paradoxical difference in cortical thickness between groups as the preservation of cortical thickness in trainees contrasted with natural, time-related thinning in controls. As evidence for this suggestion, Román and colleagues (2016) cite an investigation that observed developmentally normal cortical thinning across time in participants ranging in age from 5 – 32 years (Zhou, Lebel, Treit, Evans, & Beaulieu, 2015). However, Zhou and colleagues' (2015) study scanned participants approximately four years apart and identified adolescence, rather than young adulthood, as a period of accelerated thinning. Conversely, Román and colleagues' (2016) participants were scanned only four months apart. Observing statistically significant cortical thinning in such a short period seems biologically questionable. Therefore, Román and colleagues' (2016) supposition that their findings were due to preservation of cortical thickness in trainees relative to natural thinning in untrained controls is unlikely. Likewise, the present findings reflect a lack of support for the idea that working memory training has a significant effect on cortical thickness.

In terms of cortical surface area changes associated with working memory training, Román and colleagues (2016) described significant differences between trained and untrained groups in the right pars opercularis, with preserved surface area in the trained group and decreased area in the untrained control group. In my data, I did not observe noteworthy surface area changes across time in either training group. Again, Román and colleagues (2016) concluded that their observed decrease in pars opercularis surface area in controls related to natural, time (i.e., age) dependent declines whereas working memory training promoted preserved surface area, and therefore a relative increase associated with training compared to no training. Hogstrom and colleagues (2013) identified age associated decreases in lateral and medial areas of the prefrontal cortex in their study of natural cortical changes in 20 to 85 year olds. However, a large age range was examined (e.g., 19 years between the oldest and youngest in each age grouping considered); therefore, more change can be expected across a large time span relative to a short time span such as in my study, or that of Román and colleagues (2016). Although my findings broadly correspond with those of Román and colleagues (2016), I interpret my findings as a lack of training related change, rather than preservation of cortical features in trainees contrasted with rapid decreases in cortical metrics of controls.

One alternative explanation of why enhanced structure was not detected in working memory trainees after training is that volume, thickness, and surface area do not measure underlying neuronal characteristics that could change in response to experience (Zatorre, Fields, & Johansen-Berg, 2012). For example, synaptogenesis and pruning, changes in the morphology of dendritic spines, and vascular changes may be influencing gray matter density without significantly changing overall volumes (Zatorre et al., 2012). Wenger, Brozzoli, Lindenberger, and Lövdén's (2017) recent commentary questions the nature of expecting a perpetual increase in gray matter metrics associated with learning a new skill or task. They describe a model by which cognitive performance and regional gray matter volumes increase during the initial acquisition of a task, then while cognitive performance stabilizes with task proficiency, gray matter volume renormalizes to baseline, or near baseline levels (Wenger et al., 2017). Specifically related to gray matter volumes, they suggest initial proliferation of neuronal progenitor cells, followed by differentiation and migration, then apoptosis and maturation of select neurons (with this process being restricted to the hippocampus), and synaptic changes including dendritic branching and axon sprouting, followed by neural rewiring, stabilization of synapses, and pruning of unnecessary neural structures (Wenger et al., 2017). Based on such a model, it is likely that the brain is expanding and renormalizing before a post-intervention MRI would capture the difference. In order to resolve this issue, multiple MRI measurements would be required throughout the training process; however, such a study would require immense financial resources.

4.2 White Matter Changes Associated with Working Memory Training

A similar process involving white matter expansion and renormalization could explain our findings, and those of others, pertaining to white matter structural changes associated with working memory training. White matter volume was included in this study given that, until recently, experience dependent neural plasticity has typically excluded investigation of white matter, despite the known importance of white matter to cognition, particularly as it relates to impulse conduction and velocity of information transmission among cortical regions. Consistent with our findings pertaining to gray matter structure, white matter structure was not significantly enhanced in either training group. Furthermore, no differential white matter increases were observed between groups. Our finding is in contrast to that of Takeuchi and colleagues (2010) who identified increased myelination associated with working memory training; however, they utilized fractional anisotropy which estimates white matter fiber tract integrity, rather than volume. Their form of measurement likely provided a more specific picture of change. Furthermore, Takeuchi and colleagues (2010) compared their working memory training group to an untrained control group, rather than an active control, a methodological decision known to yield larger, though not necessarily valid, differential effects (Melby-Lervåg & Hulme, 2016).

Given that the active control participants in this study trained and improved on a processing speed task, and white matter integrity is strongly associated with speed of cognitive processing (Turken et al., 2008), perhaps I should have expected white matter increases in both our control and experimental conditions. Improvement in both groups would explain a lack of differential findings between groups. However, in this sample, neither group demonstrated notable changes in white matter volumes after training despite both groups improving on their respective training task. Wenger and colleagues' (2017) expansion and renormalization model is one explanation for these lack of findings. Specifically, according to their model, glia may proliferate and move from a resting state to an active stage, during which they increase in volume and cell surface area, then renormalize to baseline or near baseline levels (Wenger et al., 2017). As with gray matter changes, renormalization of white matter may have occurred in our sample prior to post-training scanning.

4.3 Individual Differences Associated with Structural Change after Working Memory Training

The first of my secondary hypotheses predicted that working memory training participants with lower gray and white matter volumes at baseline would exhibit larger increases in gray and white matter volumes after training. The present findings did not support this hypothesis, likely as a function of overall null effects of our training intervention. Similarly, I expected that participant age would be associated with structural change after working memory training, based on the idea that adults closer to middle age have more difficulty with the n-back task, therefore their working memory abilities would be challenged to a greater degree relative to younger participants (Jaeggi, Schmid, et al., 2008). In this sample, age was not a factor in neuroanatomical change after training. Finally, I anticipated that lower baseline estimated intelligence would be associated with greater structural change after working memory training. While this hypothesis has not previously been explored in healthy adults in a working memory training context, the hypothesis was based on previous behavioural findings, and known associations between structural integrity and cognitive functioning. Specifically, working memory trainees with higher baseline levels of cognitive performance are known to improve less, because at baseline they perform closer to ceiling levels on cognitive tests, whereas those with lower baseline estimated intelligence have greater room for improvement (Jaeggi et al., 2011). Furthermore, gray and white matter volume, particularly in frontal and parietal areas, account for a notable amount of variance in psychometric estimates of intelligence (Haier, Jung, Yeo, Head, & Alkire, 2004) and baseline differences in intelligence modulate developmental differences in cortical surface area and cortical thickness, the two metrics that give rise to cortical volume indices (Burgaleta, Johnson, Waber, Colom, & Karama, 2014; Schnack et al., 2014; Shaw et al., 2006). Therefore, I anticipated that lower cognitive reserve would be associated with greater structural change after training due to increased room for change. Findings, however, were contrary to our expectations, likely reflecting the overall null effect of working memory training in this study.

4.4 Integration of Behavioural, Functional, and Structural Correlates of Working Memory Training

As previously noted, the present study is one among a series of recent investigations into the behavioural and neural correlates of working memory training and transfer to fluid intelligence (Clark et al., 2017a, 2017b, 2017c). Results of my study fit logically into our previously published findings. First, and as expected given that the shared samples were almost identical, results of our cognitive outcomes parallel both the repeated measures ANOVAs and the Bayes factor analyses results described by Clark and colleagues (2017c). Collectively, these findings support the growing consensus that working memory training is not associated with enhanced performance on cognitive tasks, at least when compared to an active control group.

However, the notable difference between findings in other studies that utilize an active versus a passive control group suggests that some common factor of participating in a computerized training program may account for changed cognitive task performance (Melby-Lervåg & Hulme, 2016). While Melby-Lervåg and Hulme (2016) suggest this common factor to be Hawthorne and expectancy effects, Au and colleagues (2016) argue that some factor or factors, other than Hawthorne and expectancy effects, drive the observed increases in experimental and active control groups relative to untrained, passive controls. Facets of working memory and broader components of executive functioning associated with attending to stimuli and providing a motor response (e.g., sustained attention, attentional control, cognitive and motor initiation) could represent common factors that experimental and active control groups both engage throughout their training experience. Kane, Hambrick, and Conway (2005) suggest that attentional control is a critical mechanism of working memory, and explains a large component of the relationship between working memory and fluid intelligence. Certainly, both of our training programs required participants to maintain and control their attention (e.g., resist internal and external distraction) in order to complete and improve on the training games. However, in our sample, it is unlikely that similar recruitment of attentional control mediated the effects of training on cognitive performance. This is presumed because in Clark and colleagues (2017c), we did not observe significant differences in cognitive performance between the passive control group and either training group, with one exception: SDRT maintenance task performance was higher in the active control group relative to the passive control group. Therefore, within our recent series of studies, including this present study, the lack of differential change can be attributed to neither Hawthorne nor expectancy effects. Rather, our findings suggest a general lack of near or far transfer associated with working memory training on cognitive outcomes. However, this does not preclude the potential presence of an unknown factor influencing structural metrics, considering that we did not compare our structural findings to neuroanatomical estimates garnered from an untrained, passive control group. However, this is unlikely given that we did not observe robust group effects of time (i.e., training), when averaging between groups (i.e., ANOVA effects of time) or when examining mean scores and effect sizes within each group before and after training.

Despite the lack of significant training related change in near and far transfer cognitive measures, participants in both training groups improved over time in their respective training tasks. In Clark and colleagues (2017b), we reported decreased activation patterns during dual n-back task performance in participants whose training program consisted of n-back and dual n-back training (i.e., the working memory training group). This same decrease in activation was not observed in active controls who were not repeatedly exposed to n-back training. Findings suggest that working memory trainees experienced an increase in efficiency in neural networks associated with the working memory task. Such a finding is consistent with Wenger and colleagues' (2017) expansion and renormalization model, which suggests that neural structural changes that underlie neural function (e.g., synaptic expansion and pruning, glial proliferation, activation, and deactivation) leads to increased efficiency. However, collective consideration of the present data and the data from our previous three studies (Clark et al., 2017a, 2017b, 2017c) point to a general lack of significant, consistent, and meaningful change in functional, structural, or behavioural measures associated with working memory training.

4.5 Working Memory Training Does Not Work

Based on empirical evidence available at the time, and largely relating to Jaeggi and colleagues' (2008) seminal investigation that supported intelligence enhancement as a function of dual n-back working memory training, Sternberg (2008) offered two bold statements: “increasing fluid intelligence is possible after all” (pp. 6791) and “fluid intelligence is trainable to a significant and meaningful degree” (pp. 6791). In contrast, and based on the results of this study and our other recent studies (Clark et al., 2017a, 2017b, 2017c) we offer our own bold statement: working memory training does not work. However, we offer that statement with the following caveats attached: working memory training does not work, for this population, using the specific training program we tested, and the circumstances (i.e., methodologies) employed. From an empirical perspective, knowing what does not work is as beneficial as knowing what does, as it allows us to continually adjust our scientific approach, refine theories, and point resources toward making clinically relevant discoveries. In the case of working memory training, given the results of the present study, we can confidently claim that the working memory training program we used did not benefit participants over and above our active control condition. Behaviourally, we can further state that participants who completed either form of training did not benefit from training relative to passive controls who did not receive exposure to a training program (Clark et al., 2017c). Overall, these findings coincide with recent meta-analyses concluding the null cognitive effects of working memory training (Dougherty et al., 2016; Melby-Lervåg & Hulme, 2016; Melby-Lervåg et al., 2016)

4.6 Limitations

Several limitations are present in this study, each of which give rise to potential areas of further investigation. First, as with most working memory training studies and longitudinal neuroimaging studies, this study could benefit from a larger sample size. Although the sample size of 24 participants per group exceeded the recommended minimum of 20 per group (Simmons et al., 2011), and exceeded the sample size of the only other study to compare structural correlates of adaptive computerized working memory training to an active control group (i.e., Metzler-Baddeley et al., 2016), this study was underpowered based on power calculations conducted during study conceptualization which suggested a minimum of 35 participants per group. However, at that time, there were no effect sizes available from neuroanatomical studies of working memory training. Since then, Colom and colleagues' (2016) found large (d = 1.15) gray matter volumetric increases in the posterior cingulate of healthy adults after n-back working memory training. Using their effect size, and based on a power of 0.8 with $\alpha = .05$ in a 2 (group) x 2 (time) repeated measures Analysis of Variance (RM-ANOVA), calculation in G*Power yielded a required total sample size of 26 participants per group, suggesting that my study was slightly under-sampled at 24 participants per group. A post-hoc power calculation based on null findings from my study is not informative regarding statistical power given that the obtained effect size used in a post-hoc power calculation would by its nature yield low power (Hoenig & Heisey, 2001), therefore post-hoc power calculations were not performed.

Furthermore, it is possible that the regions of interest in this study were too small, or conversely, too large, to identify statistically significant changes between our groups. Regarding the former, the investigation of neuroanatomical changes associated with working memory training remains in an early stage; therefore, perhaps investigating larger regions of interest (e.g., entire lobes or structures) would have revealed significant differences that could then be followed up more specifically, thereby reducing the need for a large multiple of comparisons (e.g., as per Colom et al., 2016). Alternatively, perhaps the regions of interest I focused on were too large or otherwise imprecisely defined to identify differential changes. My use of the Desikan-Killiany atlas (Desikan et al., 2006) via the processing steps recommended by Reuter and colleagues (2012) resulted in regions of analysis that, while anatomically accurate, do not perfectly correspond to areas investigated in functional neuroimaging studies. For example, the dorsolateral prefrontal cortex is consistently implicated in functional neuroimaging investigations of working memory and fluid intelligence (Owen et al., 2005). Since the dorsolateral prefrontal cortex includes portions of the superior, middle, and inferior frontal gyri (Raz, 1997; Raz, Briggs, Marks, & Acker, 1999), changes specific to the dorsolateral prefrontal cortex may have been spread across broader anatomical regions of interest and therefore not detected in this study. However, if robust changes were present in the dorsolateral prefrontal cortex, likely there would have been at least some, even if not statistically significant, differences in superior, middle, or inferior frontal gyri metrics in working memory trainees, yet this was not the case as evidenced by within-group effects close to zero.

Another limitation of this study, and similar studies (e.g., Colom et al., 2016; Metzler-Baddeley et al., 2016; Román et al., 2016), is that the data acquisition and analysis techniques may not be sensitive enough to detect training related changes (Thomas & Baker, 2013). For example, my study was not designed to detect changes at the molecular or chemical level, which may be present as an underlying mechanism influencing structural, functional, and behavioural change in response to learned stimuli. Furthermore, this study primarily focused on gray matter and, although overall white matter volumes were analyzed, more specific measures of white matter integrity (e.g., as estimated through fractional anisotropy) were not included.

Regarding the composition of the working memory training program, we included three different training games. This decision arose from concerns expressed by the game developers that participant performance might be compromised by boredom or fatigue if they were exposed to a single game for the entire training session. The consequence of this decision was that participants were only exposed to the adaptive dual n-back aspect of training for 18 training minutes of the 24.5-minute training session, with the remainder of the session composed of visual n-back games. However, it is unlikely that this decision greatly influenced our outcomes as others have found dual n-back to be as effective as single n-back tasks for enhancing fluid intelligence (Jaeggi et al., 2010). Nevertheless, the range of adaptation was more narrow on the single n-back tasks (i.e., maximum difficulty level of 3-back) which accounted for a quarter of the overall training time. This ceiling may have limited the extent to which the Memory Match and Memory Match Overload games challenged underlying working memory processes, a theoretically necessary aspect of enhancement (Lövdén et al., 2010).

Finally, this study rests on the theoretical position relied on throughout the last decade of working memory training research, which states that by enhancing working memory and associated neural correlates, working memory and associated cognitive abilities ought to be enhanced as well (Jaeggi, et al., 2008). However, some describe this foundation as a naïve, “physical-energetic” model based on the idea that continuously challenging a limited capacity cognitive resource will lead to increased capacity, much like repeatedly using a muscle will strengthen the muscle (Melby-Lervåg & Hulme, 2013). Melby-Lervåg and Hulme (2013) argue that such a conceptualization is vague and lacks an explanation of the specific mechanisms associated with increased working memory capacity after working memory training. Although the theoretical conceptualization for this study is backed by well-established principles related to learning and neuroplasticity, we were unable to demonstrate that working memory training enhanced working memory task performance (near transfer), performance on theoretically distal tasks (far transfer), or neuroanatomical metrics to any significant degree. Identifying and examining mechanisms of action of working memory training and transfer, which will require the resolution of problems such as methodological heterogeneity and small sample sizes, may strengthen the theoretical foundation upon which behavioural (i.e., cognitive task performance) aspects of this field rest.

4.7 Future Directions

In order to resolve current theoretical challenges in the field, studies that are both well-powered and methodologically rigorous are needed to detect change, and to investigate behavioural, cognitive, or biological moderators and mediators of change associated with working memory training. However, such studies are difficult to conduct due to the immense resources necessary, hence the tendency for studies to be either well-powered *or* methodologically rigorous. In the meantime, the field could greatly benefit from a meta-analysis based on individual participant data rather than aggregate effect sizes or statistics at the group level. Individual participant data meta-analyses allow the investigation of multiple effects, including moderating and mediating effects, by analyzing a larger dataset garnered from multiple studies (Cooper & Patall, 2009) rather than relying on effect sizes from underpowered studies.

Similarly, accessing and combining shared cognitive and neuroimaging data will allow for analysis of moderators and mediators without the need to allocate resources to additional underpowered investigations. For example, using existing data (including data collected as part of this study), a future study could utilize FreeSurfer tools to manually draw and isolate cortical regions of interest that better correspond to functional neuroimaging findings. As previously discussed, the dorsolateral prefrontal cortex is consistently activated during working memory and fluid intelligence task performance, and therefore represents a logical region of interest to analyze via a manual segmentation process using already collected neuroanatomical images.

By sharing and aggregating data to strengthen the power to detect effects within existing data, resources can be allocated toward identifying underlying processes of working memory training that are potentially present yet not being identified by current measurement techniques. For example, as supported by my findings and those of others (e.g., Colom et al., 2016; Metzler-Baddeley et al., 2016; Román et al., 2016), MRI has failed to detect robust effects of working memory training on cortical surface area, thickness, and volume estimates. However, MRI is not sensitive to numerous other potential changes that occur in response to working memory training. Using Magnetic Resonance Spectroscopy, Jung and colleauges (2009) identified a positive association between gray matter N-acetylaspartate (NAA) levels and scores on performance based intelligence tasks (i.e., block design, matrix reasoning). In future studies, such alternate, non-invasive, neuroimaging techniques could be used to identify metabolite concentrations in a priori regions of interest associated with working memory training and intelligence.

Finally, whether by accessing existing data or collecting new data, investigation of white matter pathways and indices of density would fill a gap left by my study. Takeuchi and colleagues (2010) were the first to explore the influence of working memory training on white matter structural connectivity. By measuring fiber tracts using voxel-based analysis of fractional anisotropy, a metric of diffusion used to estimate fiber density and myelination, the authors identified positive associations between working memory training and white matter integrity in regions important for working memory, specifically, areas adjacent to the intraparietal sulcus and the anterior corpus callosum (Takeuchi et al., 2010). Similarly, in a recent n-back working memory training study, Salminen and colleagues (2016) detected increased white matter integrity in the anterior corpus callosum and throughout the left hemisphere in dual n-back trainees relative to both active and passive control groups. The missing piece from our recent series of investigations, including the present study, is examination of white matter structural integrity along pathways associated with working memory and fluid intelligence. Future investigation of such data, and integration with our previous structural, functional, and behavioural findings, will complete a thorough examination of the potential behavioural and neural changes associated with n-back working memory training, at least in 18-40 year old healthy, community-dwelling adults.

4.8 Clinical Relevance of Findings

My own observation of non-clinical settings, and experience within clinical settings, suggests an abundance of curiosity on the part of the lay public, and health care professionals, regarding what, if anything, an individual can do to strengthen their current cognitive resources and reduce the likelihood of future declines. Based on the present findings, our associated previous findings (Clark et al., 2017a, 2017b, 2017c), and the conclusions of others (Au et al., 2015, 2016; Dougherty et al., 2016; Karbach & Verhaeghen, 2014; Melby-Lervåg & Hulme, 2013, 2016; Melby-Lervåg et al., 2016; Schwaighofer et al., 2015; Weicker et al., 2016), the field remains in a state of flux. For the care providers who ask, or are asked, "does brain training work", or "what should I recommend for those who are concerned about preserving or enhancing their cognitive abilities?", a reasonable recommendation, based on current evidence, is that we do not yet know under what circumstances and to what extent any cognitive activity serves to enhance intelligence or stave off cognitive declines. Although positive cognitive outcomes have emerged from investigations of other forms of training, such as physical exercise (Chang, Labban, Gapin, & Etnier, 2012; Guure, Ibrahim, Adam, & Said, 2017; Sofi et al., 2011), as with working memory training, findings remain preliminary and subject to criticism. However, such a conundrum should never preclude the scientific community from continuing to investigate potential approaches to enhance the human condition. In the meantime, current recommendations appear relevant in clinical settings: exercise, hydration, adequate intake of vitamins and minerals, social activity, and cognitive activity are all behaviours actively recommended for the maintenance or improvement of brain health. However, none of these behaviours have been unequivocally demonstrated to improve intelligence or stave off age or disease related cognitive declines in healthy individuals.

4.9 Conclusion

Upon consideration of the broad debate surrounding brain training, the original consensus statement and rebuttal appear to be arguing similar albeit distinct points. While the original consensus statement focused on the exaggerated marketing practices of commercial brain training programs, the reply defends a decade of research justifying the continued exploration of cognitive training as a means to improve human performance. Similarly, those who argue that working memory training is ineffective express concern that discrete findings continue to be manipulated in favour of training related effects, despite a lack of evidence that these effects are statically or practically meaningful (Melby-Lervåg & Hulme, 2016; Moody, 2009). Nevertheless, findings based on reasonably well-designed, independent studies have led the scientific community to endorse proceeding with the investigation of the potential benefits of cognitive training, albeit with appropriate caution regarding drawing conclusions, given that the potential benefits of effective cognitive training are widespread and valuable (An open letter to the Stanford Center on Longevity, n.d.; Max Planck Institute for Human Development and Stanford Center on Longevity, 2014). Thus, despite the continued polarization within the literature, or rather, in an attempt to resolve discrepancies and identify true and meaningful effects, cognitive training remains a field worthy of continued exploration with appropriate scientific rigour.

Importantly, given the heterogeneity among different types of training interventions, one might expect variability in results. The solution to this problem lies in the ability of investigators to conduct well-designed and well-controlled studies, and make rigorous statistical choices. The solution further lies in researchers' willingness to avoid the temptation to intentionally or inadvertently cherry-pick findings or spin their results in a particular direction; rather, results ought to be presented, whether null or significant, in a clear and forthright manner. In this dissertation, I transparently presented the null findings of this study, which, commensurate with several others, fail to support the hypothesis that working memory training leads to significant and meaningful change. I suggest, therefore, that this line of research move in one of two directions: a) move meta-analytic attention to individual data meta-analyses of behavioural, structural, or neuroimaging results to resolve the pervasive issue of low power in methodologically stringent studies, or b) concentrate investigative resources on identifying and refining mechanisms of change in clinical populations who may benefit more from training interventions meant to influence broad cognitive abilities.

References

An open letter to the Stanford Center on Longevity (n.d.). Retrieved from https://www.cognitivetrainingdata.org/the-controversy-does-brain-training-work/response-letter/

Andreasen, N. C., Flaum, M., Swayze, V., O'Leary, D. S., Alliger, R., Cohen, G., … Yuh, W. T. (1993). Intelligence and brain structure in normal individuals. *The American Journal of Psychiatry*, *150*(1), 130–134. doi:10.1176/ajp.150.1.130

Ashton, M. C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment*, *91*(4), 340–345. doi:10.1080/00223890902935878

Au, J., Buschkuehl, M., Duncan, G. J., & Jaeggi, S. M. (2016). There is no convincing evidence that working memory training is NOT effective: A reply to Melby-Lervåg and Hulme (2015). *Psychonomic Bulletin and Review*, *23*(1), 331–337. doi:10.3758/s13423-015-0967-4

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: a meta-analysis. *Psychonomic Bulletin and Review*, *22*(2), 366–377. doi:10.3758/s13423-014-0699-x

Baddeley, A.D. (1992). Working Memory. *Science*, *255*(5044), 556–559. doi:10.1126/science.1736359

Baddeley, A.D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417-423. doi:10.1016/S1364-6613(00)01538-2

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829–839. doi:10.1038/nrn1201

Baddeley, A. D. (2017). Working memory: Theories, models, and controversies. In *Exploring*

Working Memory: Selected works of Alan Baddeley (pp. 333–369). London: Routledge.

Baddeley, A. D., & Hitch, G. (1974). Working Memory. *Psychology of Learning and Motivation*, *8*(5044), 47–89. doi:10.1016/S0079-7421(08)60452-1

Baddeley, A.D., & Logie, R. (1999). Working memory: The multiple component model. In A. Miyake and P. Shah (Eds.) *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). Cambridge University Press.

Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., … Willis, S. L. (2002). Effects of cognitive training interventions with older adults: A randomized controlled trial. *JAMA: The Journal of the American Medical Association*, *288*(18), 2271–2281. doi:10.1001/jama.288.18.2271.

Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex*, *49*(5), 1195–1205. doi:10.1016/j.cortex.2012.05.022

Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science*, *8*(4), 445–454. doi:10.1177/1745691613491271

Boyke, J., Driemeyer, J., Gaser, C., Buchel, C., & May, A. (2008). Training-induced brain structure changes in the dlderly. *Journal of Neuroscience*, *28*(28), 7031–7035. doi:10.1523/JNEUROSCI.0742-08.2008

Burgaleta, M., Johnson, W., Waber, D. P., Colom, R., & Karama, S. (2014). Cognitive ability changes and dynamics of cortical thickness development in healthy children and adolescents. *NeuroImage*, *84*, 810–819. doi:10.1016/j.neuroimage.2013.09.038

Burgess, G. C., Gray, J. R., Conway, A. R. A., & Braver, T. S. (2011). Neural mechanisms of interference control underlie the relationships between fluid intelligence and working memory span. *Journal of Experimental Psychology. General*, *140*(4), 674–692. doi:10.1037/a0024695.Neural

Buschkuehl, M., & Jaeggi, S. M. (2010). Improving intelligence: A literature review. *Swiss Medical Weekly, 140*(19), 266-272. doi:smw-12852

Buschkuehl, M., Jaeggi, S. M., & Jonides, J. (2012). Neuronal effects following working memory training. *Developmental Cognitive Neuroscience, 2*, S167-179. doi:10.1016/j.dcn.2011.10.001

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., Kupfer, D. J., III, C. F. R., … Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(2), 193–213. doi:10.1016/0165-1781(89)90047-4

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the raven progressive matrices test. *Psychological Review*, *97*(3), 404–431. doi:10.1037/0033-295X.97.3.404

Cattell, R. B., & Cattell, A. K. S. (1959). *Handbook for the culture fair inteligence test: A measure of "g", scale 3, forms A and B*. Champaign, IL: Institute for Personality and Ability Testing.

Chang, Y. K., Labban, J. D., Gapin, J. I., & Etnier, J. L. (2012). The effects of acute exercise on cognitive performance: A meta-analysis. *Brain Research, 1543*, 87-101. doi:10.1016/j.brainres.2012.02.068

Chooi, W.T., & Thompson, L. A. (2012). Working memory training does not improve intelligence in healthy young adults. *Intelligence*, *40*(6), 531–542. doi:10.1016/j.intell.2012.07.004

Christoff, K., Prabhakaran, V., Dorfman, J., Zhao, Z., Kroger, J. K., Holyoak, K. J., & Gabrieli, J. D. E. (2001). Rostrolateral prefrontal cortex involvement in relational integration during reasoning. *NeuroImage*, *14*(5), 1136–1149. doi:10.1006/nimg.2001.0922

Clark, C. M. (2017). *On the malleability of human cognition: Working memory training and transfer* (Doctoral Dissertation). Retrieved from https://prism.ucalgary.ca/ handle/ 11023/3642

Clark, C. M., Lawlor-Savage, L., & Goghari, V. M. (2017a). Comparing brain activations associated with working memory and fluid intelligence. *Intelligence*, *63*, 66–77. doi:10.1016/j.intell.2017.06.001

Clark, C. M., Lawlor-Savage, L., & Goghari, V. M. (2017b). Functional brain activation associated with working memory training and transfer. *Behavioural Brain Research, 334*, 34–49. doi:10.1016/j.bbr.2017.07.030

Clark, C. M., Lawlor-Savage, L., & Goghari, V. M. (2017c). Working memory training in healthy young adults: Support for the null from a randomized comparison to active and passive control groups. *PLoS ONE*, *12*(5). doi:10.1371/journal.pone.0177707

Colom, R., & García-López, O. (2003). Secular gains in fluid intelligence: Evidence from the culture-fair intelligence test. *Journal of Biosocial Science*, *35*(1), 33–39. doi:10.1017/S0021932003000336

Colom, R., Haier, R. J., Head, K., Álvarez-Linera, J., Quiroga, M. Á., Shih, P. C., & Jung, R. E.

(2009). Gray matter correlates of fluid, crystallized, and spatial intelligence: Testing the P-FIT model. *Intelligence*, *37*(2), 124–135. doi:10.1016/j.intell.2008.07.007

Colom, R., Martínez, K., Burgaleta, M., Román, F. J., García-García, D., Gunter, J. L., … Thompson, P. M. (2016). Gray matter volumetric changes with a challenging adaptive cognitive training program based on the dual n-back task. *Personality and Individual Differences*, *98*, 127–132. doi:10.1016/j.paid.2016.03.087

Colom, R., Rebollo, I., Palacios, A., Juan-Espinosa, M., & Kyllonen, P. C. (2004). Working memory is (almost) perfectly predicted by g. *Intelligence*, *32*(3), 277–296. doi:10.1016/j.intell.2003.12.002

Colom, R., Román, F. J., Abad, F. J., Shih, P. C., Privado, J., Froufe, M., … Jaeggi, S. M. (2013). Adaptive n-back training does not improve fluid intelligence at the construct level: Gains on individual tests suggest that training may enhance visuospatial processing. *Intelligence*, *41*(5), 712–727. doi:10.1016/j.intell.2013.09.002

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. . (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*(2), 163–183. doi:10.1016/S0160-2896(01)00096-4

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547–552. doi:10.1016/j.tics.2003.10.005

Cooper, H., & Patall, E. A. (2009). The relative benefits of meta-analysis conducted with individual participant data versus aggregated data. *Psychological Methods*, *14*(2), 165–176.

doi:10.1037/a0015565

Cowan, N., Elliott, E. M., Saults, S. J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51*(1), 42–100. doi:10.1016/j.cogpsych.2004.12.001

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., … Oja, P. (2003). International physical activity questionnaire: 12-Country reliability and validity. *Medicine and Science in Sports and Exercise*, *35*(8), 1381–1395. doi:10.1249/01.MSS.0000078924.61453.FB

Crook, T. H., Bartus, R. T., Ferris, S. H., Whitehouse, P., Cohen, G. D., & Gershon, S. (1986). Age-associated memory impairment: Proposed diagnostic criteria and measures of clinical change - report of a national institute of mental health work group. *Developmental Neuropsychology*, *2*(4), 261–276. doi:10.1080/87565648609540348

Curran, S. L., Andrykowski, M. A., & Studts, J. L. (1995). Short form of the Profile of Mood States (POMS-SF): Psychometric information. *Psychological Assessment*, *7*(1), 80–83. doi:10.1037/1040-3590.7.1.80

Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., … Killiany, R. J. (2006). An automated labeling system for subdeviding the human cerebral cortex on MRI scans into gryal based regions of interest. *NeuroImage, 31*(3), 968–980. doi:10.1016/j.neuroimage.2006.01.021

Dickerson, B. C., Fenstermacher, E., Salat, D. H., Wolk, D. A., Maguire, R. P., Desikan, R., … Fischl, B. (2008). Detection of cortical thickness correlates of cognitive performance:

Reliability across MRI scan sessions, scanners, and field strengths. *NeuroImage*, *39*(1), 10-18. doi:10.1016/j.neuroimage.2007.08.042

Dougherty, M. R., Hamovitz, T., & Tidwell, J. W. (2016). Reevaluating the effectiveness of n-back training on transfer through the Bayesian lens: Support for the null. *Psychonomic Bulletin and Review*, *23*(1), 306–316. doi:10.3758/s13423-015-0865-9

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004). Changes in gray matter induced by training. *Nature*, *427*(6972), 311–312. doi:10.1038/427311a

Draganski, B., Gaser, C., Kempermann, G., Kuhn, H. G., Winkler, J., Buchel, C., & May, A. (2006). Temporal and spatial dynamics of brain structure changes during extensive learning. *Journal of Neuroscience*, *26*(23), 6314–6317. doi:10.1523/JNEUROSCI.4628-05.2006

Driemeyer, J., Boyke, J., Gaser, C., Büchel, C., & May, A. (2008). Changes in gray matter induced by learning--revisited. *PloS One*, *3*(7), e2669. doi:10.1371/journal.pone.0002669

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*(10), 475–483. doi:10.1016/S0166-2236(00)01633-7

Edwards, J. D., Wadley, V. G., Myers, R. S., Roenker, D. L., Cissell, G. M., & Ball, K. K. (2002). Transfer of a speed of processing intervention to near and far cognitive functions. *Gerontology*, *48*(5), 329–340. doi:10.1159/000065259

Edwards, J. D., Wadley, V., Vance, D., Wood, K., Roenker, D., & Ball, K. (2005). The impact of speed of processing training on cognitive and everyday performance. *Aging & Mental Health*, *9*(3), 262–271. doi:10.1080/13607860412331336788

Elbert, T., & Rockstroh, B. (2004). Reorganization of human cerebral cortex: The range of changes following use and Injury. *Neuroscientist, 10*(2), 129–141. doi:10.1177/1073858403262111

Engle, R. W., Laughlin, J. E., Tuholski, S. W., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309–331. doi:10.1037/0096-3445.128.3.309

Engvig, A., Fjell, A. M., Westlye, L. T., Moberget, T., Sundseth, O., Larsen, V. A., & Walhovd, K. B. (2010). Effects of memory training on cortical thickness in the elderly. *NeuroImage*, *52*(4), 1667–1676. doi:10.1016/j.neuroimage.2010.05.041

Erdfelder, E., Faul, F., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. doi:10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. doi:10.3758/BF03193146

Fernandez, A. (2011). The business and ethics of the brain fitness boom. *Generations - Journal of the American Society on Aging*, *35*(2), 63–69. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=66815095&site=ehost-live

Fletcher, J. (2007). What is heterogeneity and is it important? *British Medical Journal*, *334*(7584), 94–96. doi:10.1136/bmj.39057.406644.68

Foroughi, C. K., Monfort, S. S., Paczynski, M., McKnight, P. E., & Greenwood, P. M. (2016). Placebo effects in cognitive training. *Proceedings of the National Academy of Sciences*, *113*(27), 7470–7474. doi:10.1073/pnas.1601243113

Frangou, S., Chitins, X., & Williams, S. C. R. (2004). Mapping IQ and gray matter density in healthy young people. *NeuroImage*, *23*(3), 800–805. doi:10.1016/j.neuroimage.2004.05.027

FreeSurfer Software Suite (Version 5.3.0) [Computer software]. Boston, MA: Martinos Center for Biomedical Imaging. Retrieved from http://surfer.nmr.mgh.harvard.edu/

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological science*, *17*(2), 172–179. Retrieved from https://pdfs.semanticscholar.org/24e2/e553e6b52d7b9b6947cb51c2fcbcc875eec4.pdf

Gautam, P. (2009, February 7). Creating label for dlpfc [Freesurfer]. Retrieved from https://mail.nmr.mgh.harvard.edu/pipermail/freesurfer/2009-February/009614.html

Glahn, D. C., Kim, J., Cohen, M. S., Poutanen, V. P., Therman, S., Bava, S., … Cannon, T. D. (2002). Maintenance and manipulation in spatial working memory: Dissociations in the prefrontal cortex. *NeuroImage, 17*(1), 201–213. doi:10.1006/nimg.2002.1161

Gläscher, J., Rudrauf, D., Colom, R., Paul, L. K., Tranel, D., Damasio, H., & Adolphs, R. (2010). Distributed neural system for general intelligence revealed by lesion. *Proceedings of the National Academy of Sciences*, *107*(10), 4705–4709. doi:10.1073/pnas.0910397107

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, *24*(1), 13–23. doi:10.1016/S0160-2896(97)90011-8

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, *6*(3), 316–322. doi:10.1038/nn1014

Greve, D. N. (2011). An absolute beginner’s guide to surface- and voxel-based morphometric analysis. *Proceedings of the International Society for Magnetic Resonance in Medicine*, *19*, 1–7. Retrieved from http://cds.ismrm.org/protected/11MProceedings/files/ISMRM2011-8410.pdf

Guure, C. B., Ibrahim, N. A., Adam, M. B., & Said, S. M. (2017). Impact of physical activity on cognitive decline, dementia, and its subtypes: Meta-analysis of prospective studies. *BioMed Research International*, *2017*, 1–13. doi:10.1155/2017/9016924

Haier, R. J., Jung, R. E., Yeo, R. A., Head, K., & Alkire, M. T. (2004). Structural brain variation and general intelligence. *NeuroImage*, *23*(1), 425–433. doi:0.1016/j.neuroimage.2004.04.025

Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sciences*, *11*(6), 236–242. doi:10.1016/j.tics.2007.04.001

Hardy, J. L., Nelson, R. A., Thomason, M. E., Sternberg, D. A., Katovich, K., Farzin, F., & Scanlon, M. (2015). Enhancing cognitive abilities with comprehensive training: A large, online, randomized, active-controlled trial. *PLoS ONE*, *10*(9), 1–17. doi:10.1371/journal.pone.0134467

Hebb, D. O. (1949). The organization of behaviour: A neurpsychological theory. New York: Wiley.

Hoenig, J.M. & Heisey, D.M. (2001). The abuse of power: The pervasive fallacy of power calculations for data analysis. *The American Statistician*, *55*(1), 19–24. Retrived from www.vims.edu/people/hoenig_jm/pubs/hoenig2.pdf

Hogstrom, L. J., Westlye, L. T., Walhovd, K. B., & Fjell, A. M. (2013). The structure of the cerebral cortex across adult life: Age-related patterns of surface area, thickness, and gyrification. *Cerebral Cortex*, *23*(11), 2521–2530. doi:10.1093/cercor/bhs231

Huntley, J. D., Gould, R. L., Liu, K., Smith, M., & Howard, R. J. (2015). Do cognitive interventions improve general cognition in dementia? A meta-analysis and meta-regression. *BMJ Open*, 2015(5), 1–12. doi:10.1136/bmjopen-2014-005247

Ilg, R., Wohlschlager, A. M., Gaser, C., Liebau, Y., Dauner, R., Woller, A., … Muhlau, M. (2008). Gray matter increase induced by practice correlates with task-specific activation: A combined functional and morphometric magnetic resonance imaging study. *Journal of Neuroscience*, *28*(16), 4210–4215. doi:10.1523/JNEUROSCI.5722-07.2008

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(19), 6829–6833. doi:10.1073/pnas.0801268105

Jaeggi, S. M., Buschkuehl, M., Jonides, J., Shah, P., Morrison, A. B., & Chein, J. M. (2011). Short-and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences*, *108*(25), 46–60. doi:10.1073/pnas.1103228108

Jaeggi, S. M., Buschkuehl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory & Cognition*, *42*(3), 464–480. doi:10.3758/s13421-013-0364-z

Jaeggi, S. M., Schmid, C., Buschkuehl, M., & Perrig, W. J. (2008). Differential age effects in load-dependent memory processing. *Aging, Neuropsychology, and Cognition*, *16*(1), 80–102. doi:10.1080/13825580802233426

Jaeggi, S. M., Studer-Luethi, B., Buschkuehl, M., Su, Y. F., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning - implications for training and transfer. *Intelligence, 38*(6), 625–635. doi:10.1016/j.intell.2010.09.001

Jaušovec, N., & Jaušovec, K. (2012). Working memory training: Improving intelligence - changing brain activity. *Brain and Cognition*, *79*(2), 96–106. doi:10.1016/j.bandc.2012.02.007

Johansen-Berg, H. (2010). Behavioural relevance of variation in white matter microstructure. *Current Opinion in Neurology, 23*(4), 351-358. doi:10.1097/WCO.0b013e32833b7631

Jolles, D. D., Van Buchem, M. A., Crone, E. A., & Rombouts, S. A. R. B. (2013). Functional brain connectivity at rest changes after working memory training. *Human Brain Mapping*, *34*(2), 396–406. doi:10.1002/hbm.21444

Jung, R. E., Gasparovic, C., Chavez, R. S., Caprihan, A., Barrow, R., & Yeo, R. A. (2009). Imaging intelligence with proton magnetic resonance spectroscopy. *Intelligence, 37*(2), 192–198. doi:10.1016/j.intell.2008.10.009

Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: Converging neuroimaging evidence. *Behavioral and Brain Sciences, (30)*2, 135-154. doi:10.1017/S0140525X07001185

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 66–71. doi:10.1037/0033-2909.131.1.66

Karbach, J., & Verhaeghen, P. (2014). Making working memory work: A meta-snalysis of executive-control and working memory training in older adults. *Psychological Science*, *25*(11), 2027–2037. doi:10.1177/0956797614548725

Katzman, R. (1993). Education and the prevalence of dementia and Alzheimer's disease.

Neurology, *43*(1), 13–20.doi:10.1212/WNL.43.1_Part_1.13

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, *14*(7), 317-324. doi:10.1016/j.tics.2010.05.002

Kroger, J. K. (2002). Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: A parametric study of relational complexity. *Cerebral Cortex*, *12*(5), 477–485. doi:10.1093/cercor/12.5.477

Kundu, B., Sutterer, D. W., Emrich, S. M., & Postle, B. R. (2013). Strengthened effective connectivity underlies transfer of working memory training to tests of short-term memory and attention. *Journal of Neuroscience*, *33*(20), 8705–8715. doi:10.1523/JNEUROSCI.5565-12.2013

Lawlor-Savage, L., & Goghari, V. M. (2016). Dual n-back working memory training in healthy adults: A randomized comparison to processing speed training. *PLoS ONE*, *11*(4), 1–17. doi:10.1371/journal.pone.0151817

Lewis, A., & Smith, D. (1993). Defining Higher Order Thinking. *Theory Into Practice*, *32*(3), 131–137. doi:10.1080/00405849309543588

Lilienthal, L., Tamez, E., Shelton, J. T., Myerson, J., & Hale, S. (2013). Dual n-back training increases the capacity of the focus of attention. *Psychonomic Bulletin and Review*, *20*(1), 135–141. doi:10.3758/s13423-012-0335-6

Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A Theoretical Framework for the Study of Adult Cognitive Plasticity. *Psychological Bulletin*, *136*(4), 659–676. doi:10.1037/a0020080

Luders, E., Narr, K. L., Thompson, P. M., & Toga, A. W. (2009). Neuroanatomical correlates of intelligence. *Intelligence*, *37*(2), 156–163. doi:10.1016/j.intell.2008.07.002

Lumos Labs Inc. (2009). Retrieved February 15, 2013, from www.lumosity.com

Mackey, A. P., Hill, S. S., Stone, S. I., & Bunge, S. A. (2011). Differential effects of reasoning and speed training in children. *Developmental Science*, *14*(3), 582–590. doi:10.1111/j.1467-7687.2010.01005.x

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences*, *97*(8), 4398–4403. doi:10.1073/pnas.070039597

Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: A structural MRI and neuropsychological analysis. *Hippocampus*, *16*(12), 1091–1101. doi:10.1002/hipo.20233

Max Planck Institute for Human Development and Stanford Center on Longevity. (2014). *A consensus on the brain training industry from the scientific community*. Retrieved from http://longevity3.stanford.edu/blog/2014/10/15/the-consensus-on-the-brain-training-industry-from-the-scientific-community/

May, A. (2011). Experience-dependent structural plasticity in the adult human brain. *Trends in Cognitive Sciences*, *15*(10), 475–482. doi:10.1016/j.tics.2011.08.002

McCabe, D. P., Henry L. Roediger III, McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, *24*(2), 222–243. doi:10.1037/a0017619

McGurk, S. R., & Mueser, K. T. (2017). Introduction to special issue on cognitive remediation. *Psychiatric Rehabilitation Journal*, *40*(1), 1–3. doi:10.1037/prj0000263

Melby-Lervag, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, *49*(2), 270–291. https://doi.org/10.1037/a0028228

Melby-Lervåg, M., & Hulme, C. (2016). There is no convincing evidence that working memory training is effective: A reply to Au et al. (2014) and Karbach and Verhaeghen (2014). *Psychonomic Bulletin and Review*, *23*(1), 324–330. doi:10.3758/s13423-015-0862-z

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer": Evidence from a meta-analytic review. *Perspectives on Psychological Science*, *11*(4), 512–534. doi:10.1177/1745691616635612

Merzenich, M. (n.d.). Does scientific evidence show bain training works? Retrieved from https://www.cognitivetrainingdata.org/the-controversy-does-brain-training-work/

Metzler-Baddeley, C., Caeyenberghs, K., Foley, S., & Jones, D. K. (2016). Task complexity and location specific changes of cortical thickness in executive and salience networks after working memory training. *NeuroImage*, *130*, 48–62. doi:10.1016/j.neuroimage.2016.01.007

Miyake, A., & Shah, P. (1999). Toward unified theories of working memory: Emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake and P. Shah (Eds.) *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 442-481). Cambridge University Press.

Moody, D. E. (2009). Can intelligence be increased by training on a task of working memory? *Intelligence*, *37*(4), 327–328. doi:10.1016/j.intell.2009.04.005

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review*, *18*(1), 46–60. doi:10.3758/s13423-010-0034-0

Mortimer, J. A., Borenstein, A. R., Gosche, K. M., & Snowdon, D. A. (2005). Very early detection of Alzheimer neuropathology and the role of brain reserve in modifying its clinical expression. In *Journal of Geriatric Psychiatry and Neurology*, *18*(4), 218–223. doi:10.1177/0891988705281869

Mouton, P. R., Martin, L. J., Calhoun, M. E., Dal Forno, G., & Price, D. L. (1998). Cognitive decline strongly correlates with cortical atrophy in Alzheimer's dementia. *Neurobiology of Aging*, *19*(5), 371–377. doi:10.1016/S0197-4580(98)00080-3

NCS Pearson Inc. (2007). *Raven's advanced progressive matrices: Evidence of reliability and validity*. Bloomington, MN: Pearson Education, Inc.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, *7*(1), 75–79. doi:10.1038/nn1165

Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S., Burns, A. S., … Ballard, C. G. (2010). Putting brain training to the test. *Nature*, *465*(7299), 775–778. doi:10.1038/nature09042

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. In *Human Brain Mapping*, *25*(1), 46-59. doi:10.1002/hbm.20131

Prabhakaran, V., Smith, J. A. L., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. E. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's Progressive Matrices Test. *Cognitive Psychology*, *33*(1), 43–63. doi:10.1006/cogp.1997.0659

Rakic, P., Ayoub, A. E., Breunig, J. J., & Dominguez, M. H. (2009). Decision by division: making cortical maps. *Trends in Neurosciences, 32*(5), 291–301. doi:10.1016/j.tins.2009.01.007

Rauchensteiner, S., Kawohl, W., Ozgurdal, S., Littmann, E., Gudlowski, Y., Witthaus, H., … Juckel, G. (2011). Test-performance after cognitive training in persons at risk mental state of schizophrenia and patients with schizophrenia. *Psychiatry Research*, *185*(3), 334–339. doi:10.1016/j.psychres.2009.09.003

Raven, J. C. (1975). *Advanced progressive matrices sets I and II: Plan and use of the scale with a report of experimental work*. London: H.K. Lewis & Co. Ltd. .

Raz, N. (1997). Selective aging of the human cerebral cortex observed in vivo: Differential vulnerability of the prefrontal gray matter. *Cerebral Cortex*, *7*(3), 268–282. doi:10.1093/cercor/7.3.268

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. *Psychology and Aging*, *14*(3), 436–444. doi:10.1037/0882-7974.14.3.436

Redick, T. S. (2015). Working memory training and interpreting interactions in intelligence interventions. *Intelligence, 50,* 14-20. doi:10.1016/j.intell.2015.01.014

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., … Engle, R. W. (2013). No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *Journal of Experimental Psychology:*

General, *142*(2), 359–379. doi:10.1037/a0029082

Reiss, A. L., Abrams, M. T., Singer, H. S., Ross, J. L., & Denckla, M. B. (1996). Brain development, gender and IQ in children: A volumetric imaging study. *Brain: A Journal of Neurology*, *119*(5), 1763–1774. doi:10.1093/brain/119.5.1763

Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage*, *61*(4), 1402–1418. doi:10.1016/j.neuroimage.2012.02.084

Román, F. J., Lewis, L. B., Chen, C. H., Karama, S., Burgaleta, M., Martínez, K., … Colom, R. (2016). Gray matter responsiveness to adaptive working memory training: A surface-based morphometry study. *Brain Structure and Function*, *221*(9), 4369–4382. doi:10.1007/s00429-015-1168-7

Roughan, L., & Hadwin, J. A. (2011). The impact of working memory training in young people with social, emotional and behavioural difficulties. *Learning and Individual Differences*, *21*(6), 759–764. doi:10.1016/j.lindif.2011.07.011

Rudebeck, S. R., Bor, D., Ormond, A., O'Reilly, J. X., & Lee, A. C. H. (2012). A potential spatial working memory training task to improve both episodic memory and fluid intelligence. *PLoS ONE*, *7*(11), 1–9. doi:10.1371/journal.pone.0050431

Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S. R., Busa, E., … Fischl, B. (2004). Thinning of the cerebral cortex in aging. *Cerebral Cortex*, *14*(7), 721–730. doi:10.1093/cercor/bhh032

Salminen, T., Mårtensson, J., Schubert, T., & Kühn, S. (2016). Increased integrity of white matter pathways after dual n-back training. *NeuroImage*, *133*, 244–250.

doi:10.1016/j.neuroimage.2016.03.028

Salminen, T., Strobach, T., & Schubert, T. (2012). On the impacts of working memory training on executive functioning. *Frontiers in Human Neuroscience*, *6(*166), 1–14. doi:10.3389/fnhum.2012.00166

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging*, *30*(4), 507–514. doi:10.1016/j.neurobiolaging.2008.09.023

Satz, P. (1993). Brain reserve capacity on symptom onset after brain injury: A formulation and review of evidence for threshold theory. *Neuropsychology*, *7*(3), 273–295. doi:10.1037/0894-4105.7.3.273

Savage, L. (2013). *Near and Far Transfer of Working Memory Training Related Gains in Healthy Adults* (Master's Thesis). Retrieved from https://prism.ucalgary.ca/handle/11023/1123

Schmidt-Wilcke, T., Rosengarth, K., Luerding, R., Bogdahn, U., & Greenlee, M. W. (2010). Distinct patterns of functional and structural neuroplasticity associated with learning Morse code. *NeuroImage*, *51*(3), 1234–1241. doi:10.1016/j.neuroimage.2010.03.042

Schnack, H. G., van Haren, N. E. M., Brouwer, R. M., Evans, A., Durston, S., Boomsma, D. I., … Hulshoff Pol, H. E. (2014). Changes in thickness and surface area of the human cortex and their relationship with intelligence. *Cerebral Cortex*, *25*(6), 1608–1617. doi:10.1093/cercor/bht357

Scholz, J., Klein, M. C., Behrens, T. E. J., & Johansen-Berg, H. (2009). Training induces changes in white-matter architecture. *Nature Neuroscience*, *12*(11), 1370–1371. doi:10.1038/nn.2412

Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist*, *50*(2), 138–166. doi:10.1080/00461520.2015.1036274

Schweizer, S., Hampshire, A., & Dalgleish, T. (2011). Extending brain-training to the affective domain: Increasing cognitive and affective executive control through emotional working memory training. *PLoS ONE*, *6*(9), 1–7. doi:10.1371/journal.pone.0024372

Seidler, R. D., Bernard, J. A., Buschkuehl, M., Jaeggi, S., Jonides, J., & Humfleet, J. (2010). *Cognitive Training as an Intervention to Improve Driving Ability in the Older Adult*. Retrieved from https://rosap.ntl.bts.gov/view/dot/18008

Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., … Giedd, J. (2006). Intellectual ability and cortical development in children and adolescents. *Nature*, *440*(7084), 676–679. doi:10.1038/nature04513

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*(4), 628–654. doi:10.1037/a0027473

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. doi:10.1177/0956797611417632

Sofi, F., Valecchi, D., Bacci, D., Abbate, R., Gensini, G. F., Casini, A., & Macchi, C. (2011). Physical activity and risk of cognitive decline: A meta-analysis of prospective studies. *Journal of Internal Medicine*, *269*(1), 107–117.doi:10.1111/j.1365-2796.2010.02281.x

Stephenson, C. L., & Halpern, D. F. (2013). Improved matrix reasoning is limited to training on tasks with a visuospatial component. *Intelligence*, *41*(5), 341–357.

doi:10.1016/j.intell.2013.05.006

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, *8*(3), 448–460. doi:10.1017/S1355617702813248

Sternberg, R. J. (2008). Increasing fluid intelligence is possible after all. *Proceedings of the National Academy of Sciences*, *105*(19), 6791–6792. doi:10.1073/pnas.0803396105

Tajik-Parvinchi, D., Wright, L., & Schachar, R. (2014). Cognitive rehabilitation for attention deficit/hyperactivity disorder (ADHD): Promises and problems. *Journal of the Canadian Academy of Child and Adolescent Psychiatry / de l'Academie Canadienne de Psychiatrie de L'enfant et de L'adolescent*, *23*(3), 207–217. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4197521/pdf/ccap-23-207.pdf

Takeuchi, H., & Kawashima, R. (2012). Effects of processing speed training on cognitive functions and neural systems. *Reviews in the Neurosciences*, *23*(3), 289–301. doi:10.1515/revneuro-2012-0035

Takeuchi, H., Sekiguchi, A., Taki, Y., Yokoyama, S., Yomogida, Y., Komuro, N., … Kawashima, R. (2010). Training of Working Memory Impacts Structural Connectivity. *Journal of Neuroscience*, *30*(9), 3297–3303. doi:10.1523/JNEUROSCI.4611-09.2010

Takeuchi, H., Taki, Y., & Kawashima, R. (2010). Effects of working memory training on cognitive functions and neural systems. *Reviews in the Neurosciences*, *21*(6), 427–449. doi:10.1515/REVNEURO.2010.21.6.427

Takeuchi, H., Taki, Y., Sassa, Y., Hashizume, H., Sekiguchi, A., Fukushima, A., & Kawashima, R. (2011). Working memory training using mental calculation impacts regional gray matter of the frontal and parietal regions. *PLoS ONE*, *6*(8). doi:10.1371/journal.pone.0023175

Thomas, C., & Baker, C. I. (2013). Teaching an adult brain new tricks: A critical review of evidence for training-dependent structural plasticity in humans. *NeuroImage*. doi:10.1016/j.neuroimage.2012.03.069

Thompson, T. W., Waskom, M. L., Garel, K. L. A., Cardenas-Iniguez, C., Reynolds, G. O., Winter, R., … Gabrieli, J. D. E. (2013). Failure of working memory training to enhance cognition or intelligence. *PLoS ONE*, *8*(5). doi:10.1371/journal.pone.0063614

Thorndike, E. L. (1932). *The fundamentals of learning.* New York, NY, US: Teachers College Bureau of Publications. https://doi.org/10.1037/10976-000

Turken, A. U., Whitfield-Gabrieli, S., Bammer, R., Baldo, J. V., Dronkers, N. F., & Gabrieli, J. D. E. (2008). Cognitive processing speed and the structure of white matter pathways: Convergent evidence from normal variation and lesion studies. *NeuroImage*, *42*(2), 1032–1044. doi:10.1016/j.neuroimage.2008.03.057

United States District Court for the Northern District of California. (2016). *Federal Trade Commission v. Lumos Labs Inc*. Retrieved from https://www.ftc.gov/system/files/documents/cases/160105lumoslabsstip.pdf

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37*(3), 498–505. doi:10.3758/BF03192720

Vance, D., Dawson, J., Wadley, V., Edwards, J., Roenker, D., Rizzo, M., & Ball, K. (2007). The accelerate study: The longitudinal effect of speed of processing training on cognitive performance of older adults. *Rehabilitation Psychology*, *52*(1), 89–96.

https://doi.org/10.1037/0090-5550.52.1.89

Vartanian, O., Jobidon, M. E., Bouak, F., Nakashima, A., Smith, I., Lam, Q., & Cheung, B. (2013). Working memory training is associated with lower prefrontal cortex activation in a divergent thinking task. *Neuroscience*, *236*, 186–194. doi:10.1016/j.neuroscience.2012.12.060

Wadley, V. G., Benz, R. L., Ball, K. K., Roenker, D. L., Edwards, J. D., & Vance, D. E. (2006). Development and evaluation of home-based speed-of-processing training for older adults. *Archives of Physical Medicine and Rehabilitation*, *87*(6), 757–763. doi:10.1016/j.apmr.2006.02.027

Waechter, R. L., Goel, V., Raymont, V., Kruger, F., & Grafman, J. (2013). Transitive inference reasoning is impaired by focal lesions in parietal cortex rather than rostrolateral prefrontal cortex. *Neuropsychologia*, *51*(3), 464–471. doi:10.1016/j.neuropsychologia.2012.11.026

Wechsler, D. (2008). *Wechsler adult intelligence scale–fourth edition technical and interpretive manual.* San Antonio, TX: Psychological Corporation.

Weicker, J., Villringer, A., & Thöne-Otto, A. (2016). Can impaired working memory functioning be improved by training? A meta-analysis with a special focus on brain injured patients. *Neuropsychology, 30*(2), 190–212. doi:10.1037/neu0000227

Wenger, E., Brozzoli, C., Lindenberger, U., & Lövdén, M. (2017). Expansion and renormalization of human brain structure during skill acquisition. *Trends in Cognitive Sciences*, *21*(12), 930–939. doi:10.1016/j.tics.2017.09.008

Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., … Wright, E. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *JAMA: The Journal of the American Medical Association*, *296*(23), 2805–2814. doi:10.1001/jama.296.23.2805

Woollett, K., & Maguire, E. A. (2011). Acquiring "the knowledge" of London's layout drives structural brain changes. *Current Biology*, *21*(24), 2109–2114. doi:10.1016/j.cub.2011.11.018

Yuan, P., Voelkle, M. C., & Raz, N. (2018). Fluid intelligence and gross structural properties of the cerebral cortex in middle-aged and older adults: A multi-occasion longitudinal study. *NeuroImage*, *172*, 21–30. doi:10.1016/j.neuroimage.2018.01.032

Zatorre, R. J., Fields, R. D., & Johansen-Berg, H. (2012). Plasticity in gray and white: Neuroimaging changes in brain structure during learning. *Nature Neuroscience*. doi:10.1038/nn.3045

Zhou, D., Lebel, C., Treit, S., Evans, A., & Beaulieu, C. (2015). Accelerated longitudinal cortical thinning in adolescence. *NeuroImage*, *104*, 138–145. doi:10.1016/j.neuroimage.2014.10.005

Zimmerman, M. E., Brickman, A. M., Paul, R. H., Grieve, S. M., Tate, D. F., Gunstad, J., … Gordon, E. (2006). The relationship between frontal gray matter volume and cognition varies across the healthy adult lifespan. *American Journal of Geriatric Psychiatry*, *14*(10), 823–833. doi:10.1097/01.JGP.0000238502.40963.ac

Zinke, K., Zeintl, M., Rose, N. S., Putzmann, J., Pydde, A., & Kliegel, M. (2014). Working memory training and transfer in older adults: Effects of age, baseline performance, and training gains. *Developmental Psychology*, *50*(1), 304–315. doi:10.1037/a0032982

APPENDIX A: Screening Questions Completed at www.braintrainingstudy.ca

* Required

First Name: *

Preferred method of contact: *

Email

Telephone

Email address:

Phone number (including area code):

What is your current age? *

Do you speak and read English fluently? *

Yes

No

Have you ever suffered from a concussion or any other form of head trauma? *

Yes

No

Not sure

Have you ever suffered from a brain fever? *

Yes

No

Not sure

Have you ever been diagnosed with a neurological or psychiatric illness? *

Yes

No

Not Sure

Prefer not to answer

Do you now, or have you in the past three months, used benzodiazepines or illicit drugs (e.g., narcotics, stimulants, depressants / sedatives, hallucinogens, cannabis)? *

Yes

No

Not sure

Prefer not to answer

Do you have any difficulties with your vision or hearing that is not corrected by glasses, contacts, aids, or implants? *

Yes

No

Not sure

Prefer not to answer

Do you currently have a cardiovascular condition or breathing problems? *

Yes

No

Not sure

Prefer not to answer

Do you currently play brain training games on the computer? *

Yes

No

Not sure

If you do, what brain training games do you play?

How often do you play these games?

Daily

A few times a week

Once a week

Less than once a week

Do you have access to high-speed internet either at home, work, or elsewhere, and can you arrange for 20-30 minutes per day, 5 days per week, of undisturbed internet use so you can train? *

Yes

No

Are you able to attend a 2.5 hour meeting at the University of Calgary (meeting times are flexible and will be made to suit your schedule)

Yes

No

Thank you for answering these questions, and for your interest in the University of Calgary Brain Training Study! A researcher will contact you to follow-up on your

responses and/or welcome you to the study. Before submitting your responses, please briefly tell us how you heard about this study.

APPENDIX B: Online Demographic Questionnaire

Sample of self-report questionnaire completed online prior to baseline appointment. The actual questionnaire can be viewed, using a dummy participant number (e.g., 9999) at:

https://ucalgarypsych.co1.qualtrics.com/SE/?SID=SV_5o86oXknKWDVM8J

Demographic Information

1. **Gender:** Male / Female

2. **Birth date:** __day__ /__month__/__year__

3. **Self Identified Ethnic Origin**: ______________________________

4. **Is English your primary language?** Yes / No
   If no, are you able to speak, read, and write fluently in English? Yes / No

5. **Current Marital Status** (circle one number):
   1 = Single (never married)
   2 = Married
   3 = Common-Law
   4 = Divorced (not remarried)
   5= Separated
   6 = Widowed
   7 = Other: _______________please briefly describe _____________

6. **Current employment status** (circle one number):
   1 = Full-time paid work outside the home for an organization
   2 = Full-time paid work for self-owned business (self-employed)
   3 = Part-time (less than 30 hours / week) paid work outside the home for an organization
   4 = Part-time (less than 30 hours / week) paid work for self-owned business (self-employed)
   5 = Not currently employed but looking for work
   6 = On temporary leave but planning to return to employment (leave type: ______________)
   7 = Full-time parent / homemaker
   8 = Retired
   9 = Other: _______________please briefly describe _____________

7. **Highest level of education completed** (circle one number):

1 = Less than grade 8
2 = Grade 8
3 = Grade 12
5 = Some college / technical school
6 = College / technical school
7 = Some university
8 = Undergraduate Degree
9 = Master's degree
10 = Ph.D.
11 = Other: _______________please briefly describe _____________

10. **Annual Income**: Average annual household income in past 5 years (circle one number):
(Gross income based on tax returns.)
1 = Under $10,000
2 = $10,000 – $20,000
3 = $20,000 – $30,000
4 = $30,000 – $50,000
5 = $50,000 – $95,000
6 = $95,000 - 150,000
7 = $150,000 and up

Health Information

11. Are you aware of any complications that occurred during your birth? (e.g., born more than 3 weeks early, birth weight less than 5.8lbs, breech position, loss of oxygen) Yes / No
If yes, describe:

12. Have you ever suffered from a concussion where you lost consciousness for more than 30 minutes? Yes / No
If yes, date of concussion: ___Month__ / __Year___
Treatment received:

13. Have you ever suffered from any other form of head trauma? Yes / No
If yes, type of trauma:
Date of trauma: ___Month__ / __Year___
Treatment received:

14. Have you ever suffered from a brain fever? Yes / No

15. Have you ever been diagnosed with a neurological illness? Yes / No
If yes, type of illness:

16. Have you ever been diagnosed with a psychiatric illness? Yes / No
If yes, type of illness:

17. Do you now, or have you in the past three months, used benzodiazepines? Yes / No

18. Do you now, or have you in the past three months, used illicit drugs (e.g., narcotics, stimulants, depressants / sedatives, hallucinogens, cannabis). Please remember that your responses are ***confidential*** and for research purposes only. Yes / No / Would rather not say

19. Do you currently have difficulties with your vision or hearing? Yes / No
    If yes, please describe:

20. Do you currently have a cardiovascular condition or breathing problems? Yes / No
    If yes, please describe:

21. Please list any current medical or psychological conditions you have been diagnosed with, the approximate date of diagnosis, and what current treatments you are receiving (if any).

Diagnosis:	Date of Diagnosis:	Current Treatment:
e.g. Hypertension	November 1997	Eprosartan, 600mg once daily

22. Please list any other medications, vitamins, dietary supplements, and herbs you are currently taking (including dosage / frequency).

Medication / Vitamin / Supplement / Herb:	Taking since:	Current Dosage / Frequency:
e.g. St. John's Wort	January 2008	One 300 mg pill twice daily

23. Please list any psychological therapies you have used in the past month, when you began using that therapy, and the frequency of use. Examples of psychological therapies: individual counseling or therapy, group therapy, couple/family counseling or therapy, hypnosis, behaviour therapy, self-help books.

Therapy:	Using since:	Frequency:
e.g. Marriage counseling	July 2011	Once per month

24. Please fill in the following table regarding the types of cognitive activities you currently participate in. Please also include any computerized cognitive training programs or exercises you have used regularly in the last 12 months. If you are not sure if a computer game you play is considered a part of a cognitive training or a mental exercise please list it anyway. Examples of these programs include Brain Age, Brain Metrix, Cogfit, Cogmed, Lumosity, PositScience.
Please ensure each entry includes times per week AND months per year AND average duration (minutes) per time.
For example: if you play board games once a week, every month, for 60 minutes and read every day for 30 minutes and go to a painting class once every two weeks, 10 months of the year, for 90 minutes per class:

Activity:	Times per week	Months per year	Average duration (minutes) per time
Board Games	1	12	60
Reading	7	12	30
Painting class	.5	10	90

25. What is your current weight in kilograms?

26. What is your current height in centimeters?

27. The next questions are to determine your handedness. Please indicate which hand you prefer to use when performing the following activities:

	Always right	Usually right	Equally likely to use either hand	Usually left	Always left
Writing	○	○	○	○	○
Throwing a ball	○	○	○	○	○
Holding a toothbrush	○	○	○	○	○
Using a spoon	○	○	○	○	○

APPENDIX C: Baseline Questionnaire Package

There are 6 questionnaires in this package. New instructions are provided for each. Some questionnaires are double-sided so please ensure you complete both sides of the page.

1. Personality Questionnaire

Please respond to the following 8 items. Be honest – there are no right or wrong answers.

1. New ideas and projects sometimes distract me from previous ones.
   - ☐ Very much like me
   - ☐ Mostly like me
   - ☐ Somewhat like me
   - ☐ Not much like me
   - ☐ Not like me at all

2. Setbacks don’t discourage me.
   - ☐ Very much like me
   - ☐ Mostly like me
   - ☐ Somewhat like me
   - ☐ Not much like me
   - ☐ Not like me at all

3. I have been obsessed with a certain idea or project for a short time but later lost interest.
   - ☐ Very much like me
   - ☐ Mostly like me
   - ☐ Somewhat like me
   - ☐ Not much like me
   - ☐ Not like me at all

4. I am a hard worker.
   - ☐ Very much like me
   - ☐ Mostly like me
   - ☐ Somewhat like me
   - ☐ Not much like me
   - ☐ Not like me at all

5. I often set a goal but later choose to pursue a different one.

- ☐ Very much like me
- ☐ Mostly like me
- ☐ Somewhat like me
- ☐ Not much like me
- ☐ Not like me at all

6. I have difficulty maintaining my focus on projects that take more than a few months to complete.

- ☐ Very much like me
- ☐ Mostly like me
- ☐ Somewhat like me
- ☐ Not much like me
- ☐ Not like me at all

7. I finish whatever I begin.

- ☐ Very much like me
- ☐ Mostly like me
- ☐ Somewhat like me
- ☐ Not much like me
- ☐ Not like me at all

8. I am diligent.

- ☐ Very much like me
- ☐ Mostly like me
- ☐ Somewhat like me
- ☐ Not much like me
- ☐ Not like me at all

2. Personality Questionnaire

Please respond to the following 18 items. Indicate how accurately each statement describes you by placing a number between 1 and 5 to the right of the statement according to the following scale:

Extremely Uncharacteristic Of Me				Extremely Characteristic Of Me
1	2	3	4	5

Be honest – there are no right or wrong answers.

1. I would prefer complex to simple problems. _______
2. I like to have the responsibility of handling a situation that requires a lot of thinking. ______
3. Thinking is not my idea of fun. _______
4. I would rather do something that requires little thought than something that is sure to challenge my thinking abilities. _______
5. I try to anticipate and avoid situations where there is likely a chance I will have to think in depth about something. ______
6. I find satisfaction in deliberating hard and for long hours. ______
7. I only think as hard as I have to. _______
8. I prefer to think about small, daily projects to long term ones. ______
9. I like tasks that require little thought once I've learned them. ______
10. The idea of relying on thought to make my way to the top appeals to me. ______
11. I really enjoy a task that involves coming up with new solutions to problems. ______
12. Learning new ways to think doesn't excite me very much. ______
13. I prefer my life to be filled with puzzles that I must solve. ______
14. The notion of thinking abstractly is appealing to me. ______
15. I would prefer a task that is intellectual, difficult, and important to one that is somewhat important, but does not require much thought. ________
16. I feel relief rather than satisfaction after completing a task that required a lot of mental effort. _____
17. It's enough for me that something gets the job done; I don't care how or why it works. ______
18. I usually end up deliberating about issues even when they do not affect me personally. ______

Pittsburgh Sleep Quality Index (PSQI)

INSTRUCTIONS: The following questions relate to your usual sleep habits during the past month only. Your answers should indicate the most accurate reply for the majority of days and nights in the **past month**. Please answer all questions.

During the past month, what time have you usually gone to bed at night?
BED TIME: _______

During the past month, how long (in minutes) has it usually taken you to fall asleep each night?
NUMBER OF MINUTES: ________

During the past month, what time have you usually gotten up in the morning?
GETTING UP TIME: __________

During the past month, how many hours of actual sleep did you get at night? (This may be different than the number of hours you spent in bed).
HOURS OF SLEEP PER NIGHT: ________

For each of the remaining questions, check the one best response. Please answer all questions.

During the past month, how often have you had trouble sleeping because you …

Cannot get to sleep within 30 minutes

not during the past month _________	less than once a week ________	once or twice a week _________	three or more times a week ________

Wake up in the middle of the night or early morning

not during the past month _________	less than once a week ________	once or twice a week _________	three or more times a week ________

Have to get up to use the bathroom

not during the past month _________	less than once a week ________	once or twice a week _________	three or more times a week ________

Cannot breathe comfortably

not during the past month _________	less than once a week ________	once or twice a week _________	three or more times a week ________

Cough or snore loudly

not during the past month _________	less than once a week ________	once or twice a week _________	three or more times a week ________

Feel too cold

not during the past month _________	less than once a week ________	once or twice a week _________	three or more times a week ________

Feel too hot

not during the past month _________	less than once a week ________	once or twice a week _________	three or more times a week ________

Had bad dreams

not during the past month _________	less than once a week ________	once or twice a week _________	three or more times a week ________

Have pain

not during the past month _________	less than once a week ________	once or twice a week _________	three or more times a week ________

Other reason(s), please describe ____

How often during the past month have you had trouble sleeping because of this?

not during the past month _________	less than once a week ________	once or twice a week _________	three or more times a week ________

During the past month, how would you rate your sleep quality overall?
Very good _____ Fairly good _____ Fairly bad _____ Very bad _____

During the past month, how often have you taken medication (prescribed or "over the counter") to help you sleep?

not during the past month _________	less than once a week ________	once or twice a week _________	three or more times a week ________

During the past month, how often have you had trouble staying awake while driving, eating meals, or engaging in social activity?

not during the past month _________	less than once a week ________	once or twice a week _________	three or more times a week ________

During the past month, how much of a problem has it been for you to keep up enough enthusiasm to get things done?

no problem at all_________	only a very slight problem ________	somewhat of a problem _________	a very big problem ________

Profile of Mood States-Short Form:

Below is a list of words that describe feelings that people have. Please read each one carefully. Then circle ONE number corresponding to the adjective phrase which best describes HOW YOU HAVE BEEN FEELING DURING THE **PAST WEEK** INCLUDING TODAY.

		Not at all ***0***	***A Little*** ***1***	***Moderately*** ***2***	***Quite a Bit*** ***3***	***Extremely*** ***4***
1.	Tense	0	1	2	3	4
2.	Angry	0	1	2	3	4
3.	Worn-out	0	1	2	3	4
4.	Unhappy	0	1	2	3	4
5.	Lively	0	1	2	3	4
6.	Confused	0	1	2	3	4
7.	Peeved	0	1	2	3	4
8.	Sad	0	1	2	3	4
9.	Active	0	1	2	3	4
10.	On edge	0	1	2	3	4
11.	Grouchy	0	1	2	3	4
12.	Blue	0	1	2	3	4
13.	Energetic	0	1	2	3	4
14.	Hopeless	0	1	2	3	4
15.	Uneasy	0	1	2	3	4
16.	Restless	0	1	2	3	4
17.	Unable to concentrate	0	1	2	3	4
18.	Fatigued	0	1	2	3	4
19.	Annoyed	0	1	2	3	4
20.	Discouraged	0	1	2	3	4
21.	Resentful	0	1	2	3	4
22.	Nervous	0	1	2	3	4
23.	Miserable	0	1	2	3	4
24.	Cheerful	0	1	2	3	4
25.	Bitter	0	1	2	3	4
26.	Exhausted	0	1	2	3	4
27.	Anxious	0	1	2	3	4
28.	Helpless	0	1	2	3	4
29.	Weary	0	1	2	3	4
30.	Bewildered	0	1	2	3	4
31.	Furious	0	1	2	3	4
32.	Full of pep	0	1	2	3	4
33.	Worthless	0	1	2	3	4
34.	Forgetful	0	1	2	3	4
35.	Vigorous	0	1	2	3	4
36.	Uncertain about things	0	1	2	3	4
37.	Bushed	0	1	2	3	4

INTERNATIONAL PHYSICAL ACTIVITY QUESTIONNAIRE

We are interested in finding out about the kinds of physical activities that people do as part of their everyday lives. The questions will ask you about the time you spent being physically active in the **last 7 days**. Please answer each question even if you do not consider yourself to be an active person. Please think about the activities you do at work, as part of your house and yard work, to get from place to place, and in your spare time for recreation, exercise or sport.

Think about all the **vigorous** activities that you did in the **last 7 days**. **Vigorous** physical activities refer to activities that take hard physical effort and make you breathe much harder than normal. Think *only* about those physical activities that you did for at least 10 minutes at a time.

1. During the **last 7 days**, on how many days did you do **vigorous** physical activities like heavy lifting, digging, aerobics, or fast bicycling?

_____ days per week

☐ No vigorous physical activities → ***Skip to question 3***

2. How much time did you usually spend doing **vigorous** physical activities on one of those days?

_____ hours per day

_____ minutes per day

☐ Don't know/Not sure

Think about all the **moderate** activities that you did in the **last 7 days**. **Moderate** activities refer to activities that take moderate physical effort and make you breathe somewhat harder than normal. Think only about those physical activities that you did for at least 10 minutes at a time.

3. During the **last 7 days**, on how many days did you do **moderate** physical activities like carrying light loads, bicycling at a regular pace, or doubles tennis? Do not include walking.

_____ days per week

☐ No moderate physical activities → ***Skip to question 5***

4. How much time did you usually spend doing **moderate** physical activities on one of those days?

   _____ **hours per day**

   _____ **minutes per day**

   ☐ Don't know/Not sure

Think about the time you spent **walking** in the **last 7 days**. This includes at work and at home, walking to travel from place to place, and any other walking that you have done solely for recreation, sport, exercise, or leisure.

5. During the **last 7 days**, on how many days did you **walk** for at least 10 minutes at a time?

   _____ **days per week**

   ☐ No walking → ***Skip to question 7***

6. How much time did you usually spend **walking** on one of those days?

   _____ **hours per day**

   _____ **minutes per day**

   ☐ Don't know/Not sure

The last question is about the time you spent **sitting** on weekdays during the **last 7 days**. Include time spent at work, at home, while doing course work and during leisure time. This may include time spent sitting at a desk, visiting friends, reading, or sitting or lying down to watch television.

7. During the **last 7 days**, how much time did you spend **sitting** on a **week day**?

   _____ **hours per day**

   _____ **minutes per day**

   ☐ Don't know/Not sure

HEXACO Personality Scale

On the following pages you will find a series of statements about you. Please read each statement and decide how much you agree or disagree with that statement. Then write your response in the space next to the statement using the following scale:

5=strongly agree **4**=agree **3**=neutral (neither agree or disagree) **2**=disagree **1**=strongly disagree

Please answer every statement, even if you are not completely sure of your response.

1		I would be quite bored by a visit to an art gallery.
2		I plan ahead and organize things, to avoid scrambling at the last minute.
3		I rarely hold a grudge, even against people who have badly wronged me.
4		I feel reasonably satisfied with myself overall.
5		I would feel afraid if I had to travel in bad weather conditions.
6		I wouldn't use flattery to get a raise or promotion at work, even if I thought it
7		I'm interested in learning about the history and politics of other countries.
8		I often push myself very hard when trying to achieve a goal.
9		People sometimes tell me that I am too critical of others.
10		I rarely express my opinions in group meetings.
11		I sometimes can't help worrying about little things.
12		If I knew that I could never get caught, I would be willing to steal a million dollars.
13		I would enjoy creating a work of art, such as a novel, a song, or a painting.
14		When working on something, I don't pay much attention to small details.
15		People sometimes tell me that I'm too stubborn.
16		I prefer jobs that involve active social interaction to those that involve working
17		When I suffer from a painful experience, I need someone to make me feel
18		Having a lot of money is not especially important to me.
19		I think that paying attention to radical ideas is a waste of time.
20		I make decisions based on the feeling of the moment rather than on careful thought.
21		People think of me as someone who has a quick temper.
22		On most days, I feel cheerful and optimistic.
23		I feel like crying when I see other people crying.
24		I think that I am entitled to more respect than the average person is.

25		If I had the opportunity, I would like to attend a classical music concert.
26		When working, I sometimes have difficulties due to being disorganized.
27		My attitude toward people who have treated me badly is “forgive and forget”.
28		I feel that I am an unpopular person.
29		When it comes to physical danger, I am very fearful.
30		If I want something from someone, I will laugh at that person's worst jokes.
31		I’ve never really enjoyed looking through an encyclopedia.
32		I do only the minimum amount of work needed to get by.
33		I tend to be lenient in judging other people.
34		In social situations, I’m usually the one who makes the first move.
35		I worry a lot less than most people do.
36		I would never accept a bribe, even if it were very large.
37		People have often told me that I have a good imagination.
38		I always try to be accurate in my work, even at the expense of time.
39		I am usually quite flexible in my opinions when people disagree with me.
40		The first thing that I always do in a new place is to make friends.
41		I can handle difficult situations without needing emotional support from anyone
42		I would get a lot of pleasure from owning expensive luxury goods.
43		I like people who have unconventional views.
44		I make a lot of mistakes because I don’t think before I act.
45		Most people tend to get angry more quickly than I do.
46		Most people are more upbeat and dynamic than I generally am.
47		I feel strong emotions when someone close to me is going away for a long time.
48		I want people to know that I am an important person of high status.
49		I don’t think of myself as the artistic or creative type.
50		People often call me a perfectionist.
51		Even when people make a lot of mistakes, I rarely say anything negative.
52		I sometimes feel that I am a worthless person.
53		Even in an emergency I wouldn’t feel like panicking.
54		I wouldn’t pretend to like someone just to get that person to do favors for me.

55		I find it boring to discuss philosophy.
56		I prefer to do whatever comes to mind, rather than stick to a plan.
57		When people tell me that I'm wrong, my first reaction is to argue with them.
58		When I'm in a group of people, I'm often the one who speaks on behalf of the
59		I remain unemotional even in situations where most people get very sentimental.
60		I'd be tempted to use counterfeit money, if I were sure I could get away with it.