2018

Evaluation of a Cognitive Training Program for Older Adults with Mild to Moderate Cognitive Decline

Kelly Bergstrom
Minnesota State University, Mankato

Evaluation of a Cognitive Training Program for Older Adults with Mild to Moderate Cognitive Decline

By

Kelly Bergstrom

A Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Arts Degree

In

Clinical Psychology

Minnesota State University, Mankato

Mankato, Minnesota

May 2018

May 1, 2018

Evaluation of a Cognitive Training Program for Older Adults with Mild to Moderate Cognitive Decline

Kelly A. Bergstrom

This thesis has been examined and approved by the following members of the student's committee.

Jeffrey A. Buchanan
Advisor

Donald Ebel
Committee Member

Karla A. Lassonde
Committee Member

Abstract

Older adults often experience varying levels of cognitive decline. Several interventions intended to help slow the effects of cognitive decline have been studied, including cognitive training. Cognitive training involves engaging individuals, typically in a group setting, in exercises that target specific cognitive domains, such as attention, perceptual speed, memory, language, and executive functioning. Literature on cognitive training provides mixed support for its efficacy. The purpose of the current study was to determine whether a manualized, in-person cognitive training program for individuals with mild to moderate cognitive decline would lead to an improvement in cognitive performance. The program targeted six cognitive domains and was administered for 12 weeks at a local residential facility for retired nuns. Participants were assessed before the cognitive training course, after the course, and at a 12-week follow-up period. The results of this study did not support the use of cognitive training for improving functioning on most domains, but participants did see improvement on some assessments intended to measure the domains of global cognitive functioning, attention/concentration, working memory, visual memory, and visual/spatial skills. However, this study had some crucial limitations, such as having a very small, homogeneous sample size and thus, definitive conclusions should not be drawn from these findings.

Introduction

The number of older adults, defined as people over the age of 65, in the United States is currently almost 50 million people (U.S. Census Bureau, 2017). By 2050, this number is expected to rise to about 86 million people (United Nations, 2017). Because the number one risk factor for cognitive decline is getting older, with an aging population comes more need for attention to cognitive changes that occur in older individuals. However, not all cognitive decline is indicative of disease. Most people are aware of neurocognitive disorder (more commonly known as "dementia,") or at least Alzheimer's disease, the most common type of dementia. What is lesser known is that there are more mild forms of cognitive decline, such age-related cognitive decline and mild cognitive impairment (MCI).

Age-Related Cognitive Decline

Age-related cognitive decline involves some subtle changes in cognitive functioning. These changes may include a noticeable decline in processing speed, episodic memory, visual construction skills, abstraction and mental flexibility, and difficulty on verbal fluency tasks, as well as on selective and divided attention tasks (Harada, Love, & Triebel, 2013). Studies have also found that older adults show lower memory acquisition (Davis, et al., 2003; Davis, et al., 2013) and are more likely to forget information after a delay (Davis et al., 2003), than younger adults.

These cognitive changes are related to changes that occur in the brain. As one ages, changes occur in the white matter of the prefrontal cortex (Nordahl et al., 2006), and in the hippocampus and basal ganglia (Inoue et al., 2001; Small et al., 2011). However, although both physical and cognitive changes occur, the correlation between these types of changes is not perfect. For example, studies have found that individuals 75 and older struggle with tasks that

involve the hippocampus compared to individuals 60-70 years old, but this effect is not seen in learning tasks that heavily involve the basal ganglia (Krishna et al., 2012; Moustafa et al., 2012).

Mild Cognitive Impairment (MCI)

MCI, which tends to be more problematic than age-related cognitive decline, is defined as cognitive decline that is greater than what is typical for a person's age and education level, but is not severe enough to interfere with activities of daily life (Gauthier et al., 2006). More specifically, an individual with MCI must score at least 1.5 standard deviations below the average for their age on diagnostic tests (e.g. Dementia Rating Scale; Jak et al., 2009). Symptoms of MCI tend to include more frequently forgetting appointment or social engagements, losing one's train of thought during conversations, feeling overwhelmed when making decisions or completing tasks that require several steps, struggling to find one's way around familiar environments, and becoming more impulsive (Mild Cognitive Impairment, n.d.).

MCI may have no apparent biological cause, but in many cases, the brain of someone with MCI has similar features to those that are common in dementia. These features may include "plaques" (abnormal clumps of beta-amyloid protein) and "tangles" (small protein clumps of tau) in the brain, Lewy bodies (clumps of another protein), small strokes, or reduced blood flow in the brain (Mild Cognitive Impairment, n.d.).

MCI is often a precursor to developing mild/major neurocognitive disorder. However, this is not an inevitable outcome. Over 50 percent of individuals with MCI progress to mild/major neurocognitive disorder, but the rest either remain stable over time or show improvements (Gauthier et al., 2006).

Mild/Major Neurocognitive Disorder

Mild/Major neurocognitive disorder, a diagnosis with more serious implications, is commonly known as "dementia". Although the Diagnostic and Statistical Manual of Mental Disorders- Fourth Edition (DSM-IV) used the term "dementia", the Diagnostic and Statistical Manual of Mental Disorders- Fifth Edition (DSM-V) has updated the term to "mild/major neurocognitive disorder". One point to note regarding this change in terminology is that, although the term generally refers to older adults who have developed dementia, the term mild/major neurocognitive disorder is a more inclusive term that includes younger adults experiencing neurocognitive impairment due to other causes, such as traumatic brain injuries and human immunodeficiency virus (HIV) infections (American Psychiatric Association, 2013, p. 591).

According to the DSM-V, major neurocognitive disorder involves showing significant cognitive decline from previous levels of functioning in one or more cognitive domains (i.e. complex attention, executive function, learning and memory, language, perceptual-motor, or social cognition). This decline is generally evidenced by both concern from the individual or a close informant and by scores on neuropsychological or other clinical assessments that signify impairment. Additionally, the cognitive deficits interfere with activities of daily living (ADLs), do not occur exclusively in the context of delirium, and are not better explained by another mental disorder (American Psychiatric Association, 2013, p. 602).

Mild neurocognitive disorder has almost the same criteria as major neurocognitive disorder, with two exceptions. First, the cognitive decline that is observed is modest, rather than significant (although neither "modest," nor "significant" are defined in a quantifiable manner in

the DSM-V). Second, in mild neurocognitive disorder, the deficits observed do not interfere with the individual's ability to complete ADLs (American Psychiatric Association, 2013, p. 602).

Examples of how major and mild neurocognitive disorder may manifest within specific cognitive domains may help illustrate the difference between the two. Within the domain of learning and memory, someone with major neurocognitive disorder may frequently repeat something they have already said multiple times during a conversation and may have trouble keeping track of a short list of items that they need to buy while shopping. An individual with mild neurocognitive disorder, however, may repeat something they have already said to someone over a few weeks (rather than within a single conversation) and may have some minor trouble recalling recent events. Within the domain of language, an individual with major neurocognitive disorder may have significant word-finding difficulties, such that they substitute vague words and phrases into a conversation, such as "that thing" and "you know what I mean" and will often substitute pronouns for names. In more severe stages, the individual may even forget names of loved ones and may begin to engage in echolalia and eventually, become mute. An individual with mild neurocognitive disorder, on the other hand, may have some word-finding difficulties, such that they substitute general terms for specific terms. They also may have some grammatical errors and may omit or incorrectly use articles (American Psychiatric Association, 2013, p. 594).

There are several subtypes of mild/major neurocognitive disorder. Most people are familiar with Alzheimer's disease, the most common type of mild/major neurocognitive disorder. This subtype accounts for an estimated 60 to 80 percent of all mild/major neurocognitive diagnoses (Alzheimer's Association, 2017). However, there are several other subtypes, such as frontotemporal lobar degeneration, Lewy body disease, vascular disease, traumatic brain injury, and substance/medication-induced (American Psychiatric Association, 2013, p. 603).

Interventions

For all levels of cognitive decline, potential interventions have been put forth to slow the process of decline. Some interventions have been efficacious in reducing symptoms of cognitive decline, such as engaging in physical activity (Yaffe, Barnes, Nevitt, Lui, & Covinsky, 2001), participating in community activities (Béland, Zunzunegui, Alvarado, Otero, & del Ser, 2005), and having a diet rich in important nutrients, such as vitamin E, polyunsaturated fats, DHA, omega 3 fatty acids, vitamin B12, folate, and antioxidants (Morris, 2012).

Other treatments for cognitive decline, however, have not been found to be efficacious. For example, gingko biloba (Snitz et al., 2009), an extract from Gingko trees native to parts of Asia, has long been believed to enhance memory. However, Snitz and colleagues (2009) demonstrated that gingko biloba does not lead to cognitive improvements compared to placebo. Other treatments intended to enhance cognition have seen mixed results in trials. In particular, cholinesterase inhibitors and glutamate blockers, two commonly-used classes of medication to treat dementia, have been found to be efficacious in slowing the cognitive effects of the disease in some studies (Aarsland et al., 2009; Emre et al., 2010; Mori, Ikeda, & Kosaka, 2012), but others have found only modest effects (Trinh, Hoblyn, Mohanty, & Yaffe, 2003) or no effects (Schneider, Dagerman, Higgins, & McShane, 2011). Anticholinergic medications have even been found to be potentially harmful to cognitive functioning (Fox et al., 2011).

Three Types of Cognitive Intervention

Several interventions have been implemented for individuals with MCI and mild/major neurocognitive disorder, and cognitive training is one line of research that may potentially help improve cognitive symptoms or slow cognitive decline. Before examining the literature on cognitive training, it is important to first distinguish cognitive training from the similar

interventions of cognitive stimulation and cognitive rehabilitation. If these three interventions were to be put on a spectrum of least to most intensive and individualized, cognitive stimulation would come first, followed by cognitive training, and then cognitive rehabilitation.

Cognitive stimulation involves engaging in group activities that are aimed at general enhancement of cognitive functioning (rather than targeting specific domains), as well as social functioning (Clare & Woods, 2004). The basic theory behind cognitive stimulation is that a lack of cognitive activity increases the rate of cognitive decline (Woods, Aguirre, Spector, & Orrell, 2012). Thus, engaging individuals in cognitively-stimulating activities is believed to mitigate or slow cognitive decline.

In a meta-analysis by Woods and colleagues (2012), the authors looked at 15 randomized controlled trials (RCTs), involving over 700 individuals with mild to moderate dementia. All studies implemented small group activities designed to stimulate thinking and memory, such as discussions of past and present events and topics of interest, word games, puzzles, music, baking, and indoor gardening. The meta-analysis found that cognitive stimulation led to improved scores on tests of memory and other cognitive functions, as well as on quality-of-life measures for individuals in mild to moderate stages of dementia, and these effects were maintained at 1-3 month follow up sessions.

Unlike the broad activities implemented in cognitive stimulation programs, cognitive training targets more specific domains, such as memory, attention, language, executive functioning, and other skills (Clare & Woods, 2004). Cognitive training is often more structured than cognitive stimulation, as there are often manuals or structured computer programs that individuals follow for a set number of sessions. The literature on the efficacy of cognitive training will be discussed thoroughly in the two sections that follow this one.

Cognitive rehabilitation is more individualized and intensive than cognitive stimulation or cognitive training. This approach involves identifying personally-relevant goals for each individual and devising strategies for addressing these. The emphasis is on improving functioning in the everyday context rather than performance on cognitive tasks (Clare & Woods, 2004). For example, if an individual struggles with certain aspects of memory such as recalling the names of family members, then the focus of treatment would be on utilizing aspects of memory that are more functional and finding ways to compensate for those that are not.

Because cognitive rehabilitation involves different goals for each person, it is difficult to measure the efficacy of cognitive rehabilitation programs. However, researchers have shown that individuals who complete cognitive rehabilitation programs report performing at their goal levels after completing the programs and report high satisfaction with the programs (Clare et al., 2010; Huckans et al., 2013). However, when research measured participants on specific cognitive domains, the patterns of improvement across domains was not found to be consistent across trials (Huckans et al., 2013). The latter finding should not be surprising, however, as not all participants completing cognitive rehabilitation programs are trying to improve on the same cognitive domains.

Cognitive Training for Age-Related Cognitive Decline

Cognitive training inventions were originally designed for individuals with age-related cognitive decline. One of the most well-known studies, the advanced cognitive training for independent and vital elderly (ACTIVE) study, was conducted by Ball and colleagues (2002). This study included 2,832 individuals with some age-related cognitive decline from six different metropolitan areas across the United States. All participants had to have a score of 22 or higher

on the Mini-Mental State Exam (MMSE), and thus, could not have experienced substantial cognitive decline.

The participants were randomized in to four groups: memory training, reasoning training, speed-of-processing training, or a control group. The memory training focused on episodic verbal memory tasks. This involved developing mnemonic strategies for remembering word lists and sequences of items, as well as for remembering the main ideas and details of stories. The reasoning training group focused on participants' abilities to solve problems that follow serial patterns, such as identifying the pattern or letter on a travel schedule. The speed group focused on visual search skills and the ability to identify and locate visual information quickly in a divided-attention format.

All groups had 10 sessions (60-75 minutes each) of their respective cognitive training program over five to six weeks. A randomized group of sixty percent of individuals also received four sessions of booster trainings 11months later. Results indicated that the three training groups all showed significant improvement in their respective cognitive domains. There was not cross over between domains (e.g. the memory training group did not show improvement on reasoning or speed-of-processing). The control group had no significant improvement on any domains. Many of the groups sustained their improvements at a two-year follow up period. Those in the speed-of-processing and reasoning groups who received booster trainings showed enhanced gains.

A 5-year follow up of the cohort in the ACTIVE study was conducted by Willis and colleagues (2006). This follow up assessment involved providing 35-month booster trainings for those who received the 11-month booster trainings and then assessing individuals on all three cognitive domains. Participants were also asked to complete a self-report inventory that assessed

their level of difficulty with completing instrumental activities of daily living (IADLs). They also completed timed IADL tasks (e.g. timing how long it takes to look up a number in the phone book). The original ACTIVE study also included these IADL components, but did not find any compelling results with regard to the treatment groups versus the control group.

Willis and colleagues found that participants in the speed-of-processing, reasoning, and memory groups all maintained gains on their targeted cognitive abilities after five years. Additionally, the reasoning group reported significantly less difficulty completing IADLs than the control group (but this effect was not found in the memory or speed-of-processing groups). Booster training provided additional effects for those in the reasoning and the speed-of-processing groups with regard to performance in their respective domains.

A 10-year follow up to the ACTIVE study showed similar results to the 5-year follow up (Rebok et al., 2014). In this second follow up, no additional booster sessions were conducted. Rebok and colleagues found that participants in the reasoning and speed-of-processing groups maintained their cognitive improvements at the 10-year follow up period, but those in the memory group did not. Additionally, participants in all three treatment groups reported less difficulty completing IADLs compared to their reports at baseline. Those in the control group were less likely to report this effect. The individuals who had previously received booster trainings continued to show enhanced effects, but similar to the 5-year follow up, these effects only applied to those in the reasoning and the speed-of-processing groups.

Cognitive Training for Individuals with Cognitive Decline

Cognitive training has also been implemented for individuals with mild cognitive impairment or dementia. One study looked at 19 individuals with mild cognitive impairment. Participants were randomized into a control group or treatment group (Rapp, Brenes, & Marsh,

2002). The treatment group completed a multi-component memory enhancement training. The components of this training involved education about memory loss, relaxation training, memory skills training, and cognitive restructuring for memory-related beliefs. The training took place in six weekly group sessions that lasted two hours each.

The findings of this study showed no differences between the treatment group or control group on memory performance, but at post-test, the treatment group had higher perceived memory ability and stronger beliefs that they had the potential to improve. The treatment group was also more likely to use mnemonics more frequently than the control group. These results were mostly consistent at a 6-month follow up testing period, but at this follow up period, the researchers also found that individuals in the training group showed better word list recall compared to those in the control group.

Another cognitive training study was conducted with participants with mild or very mild dementia (Kanaan et al., 2014). This study involved individuals completing intensive cognitive training for four to five hours per day for 10 days over a two-week period. The training sessions included both computerized and workbook training components and targeted the domains of memory, attention (sustained, divided and switching), planning, memory, and visual-spatial processing. Participants showed improved scores after the cognitive training program compared to baseline on tasks that involved working memory, sustained attention, and switching attention.

A study by Mate-Kole and colleagues (2007) examined the effects of cognitive training for six individuals with moderate to severe dementia. This study involved conducting three one-hour training sessions per week for six weeks. The sessions involved both an instructor-led, manualized treatment (Mind Aerobics) component and an adaptive computerized cognitive training (ACCT) component. The Mind Aerobics component focused on the domains of

memory, attention, cognitive flexibility, manual dexterity, and problem-solving. The ACCT portion focused on attention, visual-spatial and motor skills, problem-solving, memory, and visual discrimination. The researchers found significant improvement on measures of global cognitive functioning and short-term memory. Collateral reports by the participants' caregivers indicated that most participants were better able to complete activities of daily living (ADLs) independently and less likely to experience cognitive failures. These collateral reports also revealed behavioral changes, most notably increased awareness of the environment, being more willing to socialize and initiate socialization, increased alertness, and improved affect.

A review by Bahar-Fuchs, Clare, and Woods (2013) examined 11 RCTs for individuals with mild Alzheimer's disease (AD) or vascular dementia. The authors distinguished between studies involving cognitive training and cognitive rehabilitation, although only one of the 11 studies assessed cognitive rehabilitation. The authors calculated an effect size for the single RCT on cognitive rehabilitation and examined the quality of the ten cognitive training RCTs. These measures of quality were based on whether participants saw improvement on global screening assessments (e.g. MMSE) and performance on neuropsychological measures that target specific domains, as well as on non-cognitive outcomes (e.g. changes in mood, activities of daily living, behavior, and general health). Secondary outcomes included changes in measures of dementia severity, outcomes for family caregivers, and outcomes for disease biomarkers for the person with dementia. The researchers also assessed for potential bias with the studies, such as selection bias, attrition bias, and reporting bias.

Unfortunately, the review did not find cognitive training interventions to be efficacious for improving cognitive functioning, mood, or activities of daily living (ADLs) for individuals with mild to moderate AD or vascular dementia. The quality of most of the studies was rated as

low to moderate. However, most aspects of the cognitive rehabilitation study were rated as high quality, and the treatment group in the study showed superior effects compared to the control group on improvement of goal performance. Cognitive rehabilitation also showed promise for improving quality-of-life.

Purpose of the Current Study

The current body of literature shows that cognitive training may have potential to help participants improve on certain cognitive domains. The purpose of the current study was to evaluate whether there was a difference in scores on cognitive assessments measured before and after a 12-week Lively Mind cognitive training course. This course was delivered twice weekly for 1-hour sessions at a residential facility for older adult retired nuns. A secondary goal of this study was to determine whether any changes in cognition are sustained at a 3-month re-test period.

Based on the results of previous studies, it is hypothesized that participants will improve from pre-test to post-test on the domains of attention/concentration, memory (both visual and verbal), perceptual speed, verbal fluency, and visual-spatial abilities. It is also hypothesized that these effects will be maintained at 3 months post-intervention.

Method

Setting

All cognitive training sessions and assessments took place at a convent for retired nuns in a small Midwestern metropolitan area. All participants resided in the facility, which functioned as both an independent living and an assisted living facility, depending on the needs of each individual.

Participants

Participants included five Caucasian women, who had either confirmed or suspected mild to moderate cognitive decline. The women ranged in age from 81 to 91 years old. All participants were retired nuns and were highly educated, with one participant reporting holding a bachelor's degree and the other four holding master's degrees. Informed consent was obtained from all participants and from legal guardians, when applicable.

Potential participants were screened using the Modified Mini-Mental Status Exam (3MS). The range of scores for eligibility in the current study were between 60 and 80 on the 3MS. However, two participants who scored above this cutoff point (88 and 91), were still allowed to participate in the study at the program director's recommendation, based on their overall day-to-day functioning. Although results for all five participants will be analyzed, results for the three individuals with qualifying scores will also be considered separately.

Of those who participated in the study, two had no memory-related diagnoses, two had unspecified memory-related diagnoses, and one had a diagnosis of dementia. Of those who had unspecified memory-related diagnoses, one had comorbid depression. The participant with a diagnosis of dementia also had comorbid diagnoses of paranoia and depression.

Two participants were taking multiple medications at the time of their participation in the cognitive training course. The participant with both an unspecified memory diagnosis and depression was taking 10mg of Aricept, 20 mg of Cymbalta, and 28 mg of Namenda XR daily, as well as .5 mg of Risperdal as needed. The participant with diagnoses of dementia, paranoia, and depression was taking 60 mg of Cymbalta and 50mg of Seroquel daily, as well as 300 mg of Gabapentin three times per day.

Lively Mind™ Course

The Lively Mind™ course is one of several Mind Aerobics™ courses developed by the New England Cognitive Center (NECC). The Lively Mind™ program states that it is "designed for individuals who may have been diagnosed with mild or early-stage dementia. Decreased memory of recent events, problems in performing sequential tasks, and difficulties with mental arithmetic may be noted" (New England Cognitive Center, 2015, p. 2).

Lively Mind™ is a manualized cognitive training course designed to be delivered in 24 in-person sessions twice weekly. The program can be administered to individual participants or to groups of 10 or fewer participants. The program is delivered by a professional who is trained in the administrations of the protocols. Training involves watching a training DVD, reading through the Lively Mind™ sourcebook, familiarizing oneself with all program materials and activity-specific trainer guidelines, and reviewing the curriculum prior to each session (New England Cognitive Center, 2015, p. 10).

Along with all other Mind Aerobics™ courses, Lively Mind™ aims to improve performance on six cognitive domains- reaction time, visual/spatial abilities, attention and concentration, memory, language, and problem-solving skills (New England Cognitive Center, 2015, p. 5). In order to target these domains, Lively Mind™ includes a variety of activities that range from easy to difficult (i.e., sessions get progressively more difficult over time). Each session includes one activity from each of the six cognitive domains, and each activity lasts between 5 and 15 minutes, with each training session design to last about 60 minutes.

Assessment Materials

Modified mini-mental status exam (3MS).

As described previously, the 3MS was used to screen participants for inclusion in the Lively Mind™ course. The 3MS provides a measure of overall cognitive functioning that is generally used to assess cognitive impairment. The exam takes 20 minutes or less to administer and has a total of 100 points possible. Its content covers the broad areas of attention and concentration, memory recall (immediate and delayed), temporal orientation, spatial orientation, language, verbal fluency, verbal reasoning, and visuospatial abilities.

The 3MS has been found to have good internal consistency ($\alpha = .87$; McDowell, Kristjansson, Hill, & Hébert, 1997). It has also been found to overall be an excellent measure of correctly detecting the presence of Alzheimer's disease ($\alpha = .93$; Tombaugh, McDowell, Kristjansson, & Hubley, 1996).

Hopkins verbal learning test-revised (HVLT-R).

The HVLT-R is a measure of verbal memory. During administration, the participant is read a list of 12 words that are part of three broad semantic categories. Thus, the most possible points one can earn on each trial is 12. The full list is read to the participant and after all words are read, the participant is asked to recall as many words from the list as they can remember, in any order. This is repeated two more times to assess immediate recall and learning abilities. After 20-25 minutes, the participant is asked to recall as many words as they can from the list (in order to assess delayed recall). In order to assess verbal recognition memory, the participant is then read a list of 24 words (12 that appeared on the list and 12 that did not), and is asked to say whether each word appeared on the original list.

The HVLT-R is highly correlated with other measures of verbal memory, such as the immediate and delayed recall (r=.75 and r=.77, respectively) of the logical memory subtest of the Wechsler Memory Scale- Revised. The HVLT-R has also been found to correctly classify Alzheimer's Disease in over 90% of cases (Shapiro, Benedict, Schretlen, & Brandt, 1999).

Forward/backward digit span.

The forward and backward digit span are tests of attention/concentration as well as working memory. They taken from the "Digit Span" subtest on the Wechsler Adult Intelligence Scale- 4th edition (WAIS-IV). On the WAIS-IV, the Digit span subtest is one of two core subtests in the working memory scale.

When administering the forward digit span test, the participant is to immediately repeat back a list of numbers read by the examiner in the exact order in which they were read. The first trial contains two digits and the digits increase in clusters of two up to nine total digits (i.e. two subsequent items contain the same number of digits and the digits increase by one after two trials). The test continues until the participant gets both items in a cluster incorrect, or until all items on the test have been completed. The backward digit span is administered the same way, but the participant is expected to repeat each number sequence in the exact opposite order in which it was read. There are 16 possible point on each form (i.e. 16 possible on the forward digit span and 20 possible on the backward digit span). The forward and backward digit span have been found to be highly correlated with other validated measures of attention (Wechsler, 2008).

Trail making test part A.

The Trail Making Test (TMT) A test is a test of perceptual speed. In this test, the participant is to connect the numbers 1-25 as quickly as they can without making mistakes. The number of seconds it takes for the participants to connect the 25 numbers is recorded. Of course,

there is no minimum time required, but it generally takes participants in non-clinical samples over 30 seconds to complete (Giovagnoli, Del Pesce, Mascheroni, Simoncelli, Laiacona, & Capitani, 1996).

When the Trail-Making Test was in its infancy, part A was found to be correlated with perceptual reasoning measures and overall intelligence, and to distinguish clinical samples from non-clinical samples (Reitan, 1959). More recent analyses have also shown that completion time increases across groups for normal controls, those with mild cognitive impairment, and those with Alzheimer's disease, respectively (Ashendorf, Jefferson, O'Connor, Chaisson, Green, & Stern, 2008).

Stroop test.

The Stroop Test is a test of executive functioning (Golden & Freshwater, 2002). This test is divided into three parts. On the first part (word page), the participant is to read a page full of the words "red," "green," and "blue," which are all printed in black ink, and are presented in a random order. There are 100 words on the page (the three color words repeated over and over). The participant is given 45 seconds to read the as many words as they can, as quickly as possible, and the number of items completed is recorded.

On the second part of the test (color page), the participant is to name a page full of lines of Xs (printed as "XXXX"). All lines are printed in red, green, or blue ink. There are 100 words on the page, presented in a random order. The participant is given 45 seconds to name the as many colors as they can, and the number of items completed is recorded.

On the third part of the test (color-word page), the participant is to name the ink color of words on the page. However, the words "RED," "BLUE," and "GREEN" are written in non-congruent ink colors (e.g. the word "GREEN" may be printed in blue or red ink, but will never

appear in green ink. There are 100 words on the page, presented in a random order. The participant is given 45 seconds to name the color of ink the letters are printed for as many items as possible.

Although the participant is not penalized for making mistakes on any of the three parts of the test, if the participant does make an error, the examiner simply says, "No," and the participant is to go back and correct their error before continuing with subsequent items. Thus, making errors slows down the participant's speed of completing items.

The Stroop Test color-word test has been found to vary based on severity of Alzheimer's disease (AD) (i.e. mild or moderate), such that those with moderate AD name the colors more slowly than those with mild AD, and both AD groups name the colors more slowly than non-clinical controls (Koss, Ober, & Delis, 1984). Scores for patients with dementia on all three parts of the Stroop Test (especially the color-word test) have also found to be correlated with scores on the Blessed Dementia Scale (BDS), a validated scale indicating severity of dementia, such that normal controls differed from those with dementia (Fisher, Freed, & Corkin, 1990).

Controlled oral word association test (COWAT).

The COWAT is a test of language and verbal fluency. In this test, participants are given a letter of the alphabet and then given 60 seconds to name as many words as they can that begin with that letter. The only stipulations are that the participants cannot list words that are proper nouns (e.g. cannot say "Bob" or "Boston" for the letter B) and cannot list the same root words multiple times with different endings (e.g. "bed, beds, and bedding" would only count as one item). This process is repeated again with another letter of the alphabet, and after both trials, the number of words the participant was able to generate for two letters of the alphabet is recorded as the score. Thus, the minimum possible score for the COWAT is zero (if no words beginning

with the specified letters are generated), and there is no maximum score. The COWAT has been found to have excellent interrater reliability and test-retest reliability (Ross, Calhoun, Cox, Wenner, Kono, & Pleasant, 2007). Additionally, the total words generated is found to correlated with other neuropsychological tests, such as the WAIS-III Vocabulary subtest, the WAIS-III Letter-Number Sequencing subtest, and the Stroop Test (Ross, Calhoun, Cox, Wenner, Kono, & Pleasant, 2007).

Brief visuospatial memory test-revised (BVMT-R).

The BVMT-R is a test of visual memory. In this test, the participant is presented with a page of six arbitrary figures, organized in two columns and three rows. The participant is asked to study the page for 10 seconds, after which time the page is removed and the participant is asked to replicate as many figures as they can recall in the exact location that they appeared on the page. This process is repeated three times, and then the participant is told that they may be asked to remember the figures later. After 20-25 minutes, the participant is asked to again replicate as many figures as they can recall in the precise location that they appear on the original page. After the delayed recall, a recognition portion of the test takes place. On this part of the test, the participant is presented 12 figures, one at a time. Six of these figures were presented originally and six were not. The participant is asked to identify whether each picture appeared on the original page.

The scoring of the BMVT-R assigns 0-2 points for each figure. No points are assigned if a figure is neither correctly drawn nor accurately placed, one point if either of those contingencies are in place, and two points if the figure is both accurately drawn and correctly placed. Thus, on each trial, 12 total point are possible, and the highest possible immediate recall

score (all three trials summed), 36 points are possible. On the delayed recall portion of the test, 12 points are possible, and on the recognition section, 12 points are also possible.

Form 1 of the BVMT, which was the form used in the current study, has been found to have an interrater reliability (.97; Benedict, Schretlen, Groninger, Dobraski, & Shpritz, 1996) and high test-retest reliability (ranging from .70 to .95, depending on the trial; Benedict et al., 1996).

Brief test of attention (BTA).

The BTA is a test of complex attention. During the administration of this test, the examiner plays a CD of a female voice reading a list of letters and numbers. The participant is to mentally keep track (counting on fingers is not allowed) of either how many numbers (in part one) or letters (in part two) were read in each string. The first letter-number string is four characters long; the second string is six characters long, and subsequent trials increase by three characters every two trials. This pattern is the same in both parts of the test. There are 10 total items on each part of the test and thus, the "Numbers" score ranges between 0 and 10, as does the "Letters" score. The BTA has been found to have a reliability ranging from .82 to .91, have no practice effects, and be strongly correlated with other tests for attention (Schretlen, Bobholz, & Brandt, 1996).

Visual puzzles.

Visual Puzzles is a test of visuospatial skills. It is taken from the perceptual reasoning scale on the WAIS-IV. For each item, the participant is asked to identify which three of six pieces displayed on the page would fit together to make an image presented at the top of the page. The participant often needs to mentally rotate items to make them fit.

The test ends when the participant has either gotten three consecutive items incorrect or has finished all items on the test. The administration of Visual Puzzles in the current study differed from administration on the WAIS-IV in two ways. First, all participants began with item 1, whereas traditionally, most participants begin with a more difficult item and earn retroactive points for the first few items, (unless they are unable to answer the starting item, at which point, the examiner administers previous items). Also, traditionally, participants are only given 30 seconds to respond and in the current study, participants were given unlimited time to complete each item. There are 26 items on Visual Puzzles and each item is assigned 1 point, yielding a total possible score of 26.

Frequency of forgetting scale.

The Frequency of Forgetting Scale (Zelinski & Gilewski, 2004) is a measure of memory self-efficacy. This is a self-report measure that contains 10 items related to memory problems. The participant is asked to rate each item on a scale of 1-7, with "1" indicating the worst problems with memory, and "7" indicating the most modest problems with memory. The items on the survey are related to general memory, and more specific items that one may have trouble with, such as names, faces, and directions. Raw scores on the Frequency of Forgetting Scale range from 10 to 70. Raw scores are converted into Rasch scores, which range from -5.14 (raw score of 10) to 5.04 (raw score of 70).

Patient health questionnaire-9 (PHQ-9).

The PHQ-9 is a measure of common symptoms of depression. This is a 10-item self-report questionnaire. Each item relates to symptoms of depression listed in the DSM-5. Each item receives a score of 0-3. If the participant denies a specific symptom, it is automatically given a score of "0". However, if they endorse a symptom, they are asked to state how many

days they have experienced that particular symptom in the past two week. Scores are assigned based on the frequency of symptoms in the following way: 1 day- 0; 2-6 days-1; 7-11 days-2; 12-14 days- 3. Thus, the overall score on the PHQ ranges from 0-30, with 0 indicating no symptoms of depression and 30 indicating several persistent symptoms of depression over the past two weeks. Indication of depressive symptoms based on PHQ-9 score is generally qualified in the following way: 5 to 9- mild depression; 10 to 14- moderate depression; 15 to 19- moderately severe depression; 20 to 27- severe depression (Kroenke & Spitzer, 2002).

The PHQ-9 has been found to have high internal consistency (α=0.89; Kroenke et al., 2001). For individuals without cognitive impairment, the PHQ-9 has been found to have a with a sensitivity of 85% and specificity of 89% for scores indicating moderate depression and above. For individual with cognitive impairment, sensitivity was found to be 89% and specificity was 71% (Boyle et al., 2011).

Procedure and Research Design

Participants were recruited by the activities director at the facility and the Lively Mind™ course was also carried out by the activities director.

This study used a quasi-experimental design in which all participants in the study completed a battery of cognitive assessments at three time periods: in the week prior to beginning the Lively Mind™ course, in the week immediately following the course, and at a 12-week follow-up period. In order to decrease the possibility of fatigue, assessments took place in three separate sessions on different days. First, participants would come in and complete the 3MS in a 20-minute scheduled session. They would later come in two more times for 30 to 45-minute session. The 3MS sessions were all completed by a faculty member at Minnesota State University, Mankato and all subsequent sessions were completed by trained graduate students.

Results

Analyses

Cohen's d effect size was used to determine whether differences were observed on measures from pre- to post-treatment. Effect sizes were used because data were only collected from five participants, which does not provide enough statistical power to use inferential statistics such as a repeated measures analysis of variance. Because data were collected at pre-test, post-test, and follow up, two effect sizes for each assessment score will be presented. Based on recommendations of Cohen (1988), the following values are used to quantify Cohen's d effect sizes: 0.00 to 0.19-no effect; 0.20 to 0.49- small effect; 0.50 to 0.79- medium effect; 0.80 or greater- large effect.

Baseline

As stated previously, two participants included in this study were above the cut-off of 80 on the 3MS and thus, did not technically qualify for the study. For exploratory purposes, it is worthwhile to consider the scores of all participants, but those who had 3MS scores within the 60-80 range will also be considered independently.

A Cohen's d calculator intended for comparing independent groups with different sample sizes was used to determine whether any baseline differences existed between participants who had 3MS scores within the range of qualifying for the Lively Mind™ group (scores between 60-80) and those outside of the range (all non-qualifying participants had scores above 80).

As can be seen in Table 1, most assessments showed large effect sizes, indicating that there was a large baseline difference on several domains between those who qualified for the Lively Mind™ course and those who did not. However, participants showed a small effect or no effect on some assessments intended to measure attention (BTA and Forward/Backward Digit

Span), perceptual speed (Trails A), visual-spatial skills (Visual Puzzles), and depressive symptoms (PHQ-9).

A Cohen's d calculator intended for measuring effect sizes in repeated-measures designs was used for calculating the effect sizes for the scores from pre-test to post-test and then from post-test to follow up. The results for all participants can be seen in Tables 2 and 3 and the results for the three participants who qualified for the Lively Mind™ course can be seen in Tables 4 and 5.

Global Cognitive Ability

As can be seen in Table 2, participants' 3MS scores had a large effect when considering all individuals in this study (0.99). However, as can be seen in Table 3, at the 3-month follow up period, participants' 3MS scores showed a large negative effect (-1.08), such that their scores essentially returned to baseline levels. As indicated in Table 4, effect sizes of 3MS scores for those who qualified for the Lively Mind™ course followed the same pattern as that for all participants, such that a large effect was seen after completion of the course (1.67) and, as indicated in Table 5, a large negative effect was seen at the 3-month follow up period (-0.99).

Verbal Memory

HVLT-R total recall.

The HVLT-R total recall score represents the total number of words a participant was able to immediately recall from the word list across all three trials. As indicated in Table 2, there was no effect observed when considering all individuals in this study (0.09). As indicated in Table 3, at the 3-month follow up period, participants' scores showed a small effect (0.23). As indicated in Table 4, those who qualified for the Lively Mind™ course saw a small effect after

completion of the course (0.27), and, as indicated in Table 5, a large effect was seen at the 3-month follow up period (1.05).

HVLT-R delayed recall.

The HVLT-R delayed recall score represents how many words participants were able to recall from the original word list after a 20-25 minute delay. As indicated in Table 2, there was a small negative effect when considering all individuals in this study (-0.34). As indicated in Table 3, at the 3-month follow up period, participants' scores showed an intermediate negative effect (-0.72). Unfortunately, effect sizes comparing scores from pre-test to post-test could not be calculated for those who qualified for the Lively Mind™ course. This can happen with in cases where there are few participants if the correlation between the two sets of scores is 1.00. As indicated in Table 5, a small negative effect was seen at the 3-month follow up period (-0.41).

HVLT-R recognition.

The HVLT-R recognition section assesses how many items that were on the original word list the participant is able to identify. The recognition score represents the difference of the participant's true positives (identifying a word as being from the list that was on the list) and false positives (identifying a word as being from the list that was not on the list) on this portion of the HVLT-R. As indicated in Table 2, there was no effect observed when considering all individuals in this study (-0.16). As indicated in Table 3, at the 3-month follow up period, participants' scores showed a small effect (0.40). As indicated in Table 4, those who qualified for the Lively Mind™ course saw a small negative effect after completion of the course (-0.20), and, as indicated in Table 5, a large effect was seen at the 3-month follow up period (2.31).

Attention

Forward digit span.

As indicated in Table 2, there was a small negative effect observed when considering all individuals in this study (-0.44). As indicated in Table 3, at the 3-month follow up period, participants' scores showed an intermediate effect (0.76). As indicated in Table 4, those who qualified for the Lively Mind™ course saw an intermediate negative effect after completion of the course (-0.64), and, as indicated in Table 5, a large effect was seen at the 3-month follow up period (2.29).

Backward digit span.

As indicated in Table 2, there was a small effect observed when considering all individuals in this study (0.41). As indicated in Table 3, at the 3-month follow up period, participants' scores showed no effect (0.00). As indicated in Table 4, those who qualified for the Lively Mind™ course saw no effect after completion of the course (0.00), and, as indicated in Table 5, a large effect was seen at the 3-month follow up period (1.32).

Brief test of attention (BTA).

For the purposes of this study, the BTA score represents the sum of the "Letters" and "Numbers" sections of the BTA. As indicated in Table 2, there was a large negative effect observed when considering all individuals in this study (-8.18). As indicated in Table 3, at the 3-month follow up period, participants' scores showed a large effect (1.48). As indicated in Table 4, those who qualified for the Lively Mind™ course followed the same pattern as that for all participants, such that a large negative effect was seen after completion of the course (-5.36), and, as indicated in Table 5, a large effect was seen at the 3-month follow up period (10.22).

Perceptual Speed and Executive Functioning

Trail-making test part A.

The score on the trail-making test is simply the number of seconds it took the participant to complete the test. Thus, a higher score indicates a slower perceptual speed. As indicated in Table 2, there was a large effect observed when considering all individuals in this study (1.69). However, this effect indicates a decline in functioning, as a higher time to complete the task indicated slower perceptual speed. As indicated in Table 3, at the 3-month follow up period, participants' scores showed a large negative effect (-0.86), which is indicative of improvement. As indicated in Table 4, those who qualified for the Lively Mind™ course saw a large effect after completion of the course (1.41), which is indicative of a decline in functioning. As indicated in Table 5, a large negative effect was seen at the 3-month follow up period (-1.18), which is indicative of improvement.

Stroop word.

As indicated in Table 2, there was no effect observed when considering all individuals in this study (0.13). As indicated in Table 3, at the 3-month follow up period, participants' scores showed a small negative effect (-0.35). As indicated in Table 4, those who qualified for the Lively Mind™ course saw an intermediate negative effect after completion of the course (-0.68), and, as indicated in Table 5, an intermediate negative effect was seen at the 3-month follow up period (-0.53).

Stroop color.

As indicated in Table 2, there was no effect observed when considering all individuals in this study (-0.08). As indicated in Table 3, at the 3-month follow up period, participants' scores showed a large negative effect (-1.21). As indicated in Table 4, those who qualified for the

Lively Mind™ course saw no effect after completion of the course (0.13), and, as indicated in Table 5, a large negative effect was seen at the 3-month follow up period (-0.99).

Stroop color-word.

As indicated in Table 2, there was a large negative effect observed when considering all individuals in this study (-2.78). As indicated in Table 3, at the 3-month follow up period, participants' scores showed a small effect (0.47). As indicated in Table 4, those who qualified for the Lively Mind™ course saw a large negative effect after completion of the course (-2.56), and, as indicated in Table 5, a small effect was seen at the 3-month follow up period (0.23).

Language

The language score in this study is the sum of the number of words generated on the COWAT over two trials. As indicated in Table 2, scores on the COWAT had a large negative effect when considering all individuals in this study (-4.54). As indicated in Table 3, at the 3-month follow up period, participants' COWAT scores showed no effect (0.00). As indicated in Table 4, those who qualified for the Lively Mind™ course saw a large negative effect after completion of the course (-8.54), and, as indicated in Table 5, a large effect was seen at the 3-month follow up period (3.33).

Visual Memory

BVMT-R total recall.

The BVMT-R total recall score is the score a participant obtained based on replicating the correct figures in their correct location on the page immediately after presentation of the sheet across three trials. As indicated in Table 2, there was a large effect observed when considering all individuals in this study (1.07). As indicated in Table 3, at the 3-month follow up period, participants' scores showed an intermediate effect (0.53). As indicated in Table 4, those who

qualified for the Lively Mind™ course saw no effect after completion of the course (0.00), and, as indicated in Table 5, a large negative effect was seen at the 3-month follow up period (-1.81).

BVMT-R delayed recall.

The BVMT-R delayed recall score indicates the participant's score based on replicating the correct figures in their correct location after a 20-25 minute delay. As indicated in Table 2, there was no effect observed when considering all individuals in this study (-0.15). As indicated in Table 3, at the 3-month follow up period, participants' scores showed a small negative effect (-0.39). As indicated in Table 4, those who qualified for the Lively Mind™ course saw a large negative effect after completion of the course (-1.00), and, as indicated in Table 5, no effect was seen at the 3-month follow up period (-0.17).

BVMT-R recognition.

The BVMT-R recognition section assesses how many figures from the original page the participant is able to identify. The recognition score represents the difference of the participant's true positives and false positives on this portion of the BVMT-R. As indicated in Table 2, there was no effect observed when considering all individuals in this study (-0.10). As indicated in Table 3, at the 3-month follow up period, participants' scores showed a large effect (0.94). As indicated in Table 4, those who qualified for the Lively Mind™ course saw no effect after completion of the course (0.00), and, as indicated in Table 5, a large effect was seen at the 3-month follow up period (1.72).

Visual-Spatial Skills

As indicated in Table 2, scores on Visual Puzzles had a small effect when considering all individuals in this study (0.20). As indicated in Table 3, at the 3-month follow up period, participants' scores showed a large effect (1.41). As indicated in Table 4, an effect size on Visual

Puzzles could not be calculated for those who qualified for the Lively Mind™ course. As indicated in Table 5, a large effect was seen at the 3-month follow up period (1.53).

Memory Self-Efficacy

Memory self-efficacy is indicated by scores on the Frequency of Forgetting scale. Lower scores on this assessment indicate more problems with memory. As indicated in Table 2, scores showed a small negative effect when considering all individuals in this study (-0.38). This effect indicates an increase in memory self-efficacy. As indicated in Table 3, at the 3-month follow up period, participants' scores showed an intermediate effect (0.60). This effect indicates a decline in memory self-efficacy. As indicated in Table 4, those who qualified for the Lively Mind™ course saw no effect after completion of the course (-0.08), and, as indicated in Table 5, a small effect was seen at the 3-month follow up period (0.35), which indicates a decline in memory self-efficacy.

Depressive Symptoms

Depressive symptoms are measured by scores on the PHQ-9. Higher scores on this assessment indicate an increase in depressive symptoms. As indicated in Table 2, scores on the PHQ-9 showed a small effect when considering all individuals in this study (0.40). This indicated that depressive symptoms increased slightly. As indicated in Table 3, at the 3-month follow up period, participants' scores showed no effect (0.11). As indicated in Table 4, those who qualified for the Lively Mind™ course saw an intermediate effect after completion of the course (0.59), indicating that depressive symptoms increased after the Lively Mind™ course. As indicated in Table 5, an intermediate negative effect was seen at the 3-month follow up period (-0.59), indicating a decrease in depressive symptoms.

Discussion

Summary of Results

In summary, the most common outcome from the pre-test to post-test period was decline and the most common outcome from post-test to follow up was improvement. Assessment scores that followed this specific pattern were those on the Forward Digit Span, BTA, Trails A, Stroop Test Color-Word, and Frequency of Forgetting when considering all participants, and those on the HVLT-R (recognition score only), Forward Digit Span, BTA, Stroop Test Color-Word, COWAT, and PHQ-9 when considering participants who qualified for the study. This general pattern is, of course, the opposite of what one would expect to find concerning an intervention targeting cognitive decline.

However, not all results in this study indicated decline. When considering all participants, some of the subtest scores intended to measure global cognitive functioning (3MS scores), attention/concentration, working memory (Backward Digit Span scores), visual memory (BVMT-R total recall scores), and visual-spatial skills (Visual Puzzles scores) showed some improvement after the Lively Mind™ course. Of those domains, global cognitive functioning is the only domain on which participants did not maintain their improvements. In fact, 3MS scores showed a large negative effect, such that the mean scores returned almost to baseline levels. Backward Digit Span scores remained stable from post-test to follow up. The scores for BVMT-R total recall and Visual Puzzles continued to show improvement from post-test to follow up.

When considering only those who qualified for the Lively Mind™ course, participants improved on assessments intended to measure the domains of global cognitive functioning (3MS) and verbal memory (HVLT-R total recall). The 3MS scores showed a large negative

effect from post-test to follow up, but the mean score remained a few points above baseline levels. The HVLT-R total recall scores continued to show improvement.

When considering all participants in this study, the improvements on the domains of attention/concentration, visual memory, and visual-spatial skills were consistent with the study's hypothesis. When considering those who qualified for the Lively Mind™ course, only the improvement of verbal memory was consistent with the hypothesis. Unfortunately, these overall results did not provide robust evidence that a Lively Mind™ cognitive training program for older adults with mild to moderate cognitive decline is efficacious. However, it is not my intention to draw any definitive conclusions regarding the efficacy of the Lively Mind™ cognitive training program, or of cognitive training in general. Several results from this study were puzzling, such as the fact that participants appeared to show some spontaneous improvements or declines from post-test to follow up testing even though no intervention was in place during this time. This finding, along with some exceptionally large effect sizes seen on some tests, led me to believe that effects seen in this study may have been a result of having a small sample size, rather than true effects of the intervention.

Overall, the current body of literature provides mixed support for the efficacy of cognitive training. Although some studies have provided support for cognitive training, the current study is more consistent with results from studies that did not find positive results and those that found mixed results. For example, Rapp, Brenes, and Marsh (2002) did not find that a multicomponent memory-enhancement training leads to improvement in actual memory performance. (The only finding was an improvement on memory self-efficacy.) Additionally, other studies have showed mixed results regarding the efficacy of cognitive training. For example, Kanaan and colleagues (2014) found that cognitive training led to improvements on the

domains of working memory, sustained attention, and switching attention, but were unable to provide evidence that improvement occurred on other domains, such as perceptual speed, category and letter fluency, and divided attention.

Sample Size and Effect Size

The low sample size available in this study, as well as some exceptionally high effect sizes, seem to suggest that the scores in this study may have been influenced by outliers. With a small sample size, effect sizes are more susceptible to influence by outliers than larger samples would be. Equations for effect size use the standard deviations in the denominator (Cohen, 1988), meaning that large standard deviations would produce smaller outcomes. This is important because having fewer participants in a study with dissimilar scores is likely to produce large standard deviations and thus, smaller effect sizes. The opposite is also possible for studies with very few participants; if there are not many participants and participants' scores are nearly identical, the standard deviation will be exceptionally small, which has the potential to produce large effect sizes.

Other factors that may influence effect size include correcting for bias, having a restricted range of potential scores, and using a non-normal distribution (Coe, 2002). In this study, restricted range and non-normal distributions may have been present. Although none of the assessments used in this study had overly restricted ranges, using participants with a specific level of cognitive functioning (i.e. mild to moderate cognitive impairment) may have solicited a restriction in the ranges uses in this study. An example of this restriction is as follows: although scores on the BTA could be between 0 and 20, four of the five participants in this study scored a 9 or 10 on the pre-test, and the last participant scored 13. (The participant who scored 13 is one who met inclusion criteria for this study.) It should be noted that although two participants

scored higher than the 3MS score required for participation in this study, this restriction may still apply, as those individuals were judged by their activities director to have a functioning level similar to those who did qualify for the study. Non-normal distribution may have also played a role in this sample, as the sample only included five individuals. In general, 30 or more participants are required to make up a normal distribution (Mordkoff, 2000).

As stated previously, exceptionally high effect sizes were seen in several of the measures. More specifically, score on the Stroop Test Color-Word, BTA, and COWAT showed exceptionally high negative effect sizes, both when considering all participants and when considering only the three participants who qualified for the study. Additionally, scores on the BTA and the COWAT showed large positive effect sizes from post-test to follow up for the individuals who qualified for the study. As can be seen in Table 6, the large negative effect size (-8.18) seen across all participants' BTA scores is accounted for by four of the five participants decreasing their score by two or three points at post-test. The large negative effect seen across all participants' Stroop Test Color-Word scores (-2.78 for all participants from pre- to post-test; -2.56 for qualifying participants) was accounted for by four of the five participants completing between 4 and 8 fewer items on the post-test than they did on the pre-test. The large negative effect seen across all participants' COWAT scores (-4.54 for all participants from pre- to post-test; -8.54 for qualifying participants) were accounted for by four of the five participants generating between 10 and 13 fewer words on the post-test than they did on the pre-test. Similar patterns to this one also explain exceptionally large positive effects seen from post-test to follow up.

Other Potential Explanations

Besides a small effect size, other explanations for the results observed should be considered. There is a possibility that the Lively Mind™ course may have been more harmful than helpful for participants. However, another recent thesis found the Lively Mind™ course to produce results that showed mainly stability or improvement from pre-test to post-test (Stypulkowski, 2017). It may have been the case that the cohort of individuals in the current study did not enjoy the course and thus, were not engaged during activities.

Another potential explanation is that a cohort effect occurred shortly before testing. For example, because all participants lived in the same facility, a negative event, such as the death of another resident, may have occurred around the time of one of the testing days, leading to several residents to have a difficult time with testing. Alternatively, a positive event may have led to increased morale, which caused participants to test exceptionally well on a specific day.

Limitations and Future Directions

The current study had several limitations. The most salient of these is that the sample size was very small (N=5). As discussed previously, a sample size this small makes achieving a normal distribution difficult and leaves room for one or two outliers to skew the effect sizes. Additionally, of the five participants in the study, there were only three participants in the study that qualified for the study. This was the consequence of having a lack of interested participants for the study. The activities director recommended the two participants who were found to have 3MS scores above 80. Because the activities director felt that their level of functioning was similar to those who qualified for the study, they were allowed to take the Lively Mind™ course and data were collected on their scores.

In addition to having a small sample size, the current study also had a homogenous group of individuals. All participants were white, female, and retired nuns. They were also all between the ages of 81 and 91. This makes it difficult to generalize findings to the entire population of older adults experiencing mild to moderate cognitive decline, as findings may apply only to individuals who share these traits.

Another limitation of this study was that there was no control group or randomization of participants taking the Lively Mind™ course. Without a control group, it cannot concluded that the cognitive training intervention was responsible for the changes observed on the outcome measures. A control group would have been especially helpful in this study, as there were several domains on which participants showed decline from pre-test to post-test and several domains on which participants showed increase, in the absence of an intervention, from post-test to follow up. Overall, a control group would allow researchers to differentiate between changes that occurred due to the cognitive training program and changes that naturally occur over time in this population. Participants were also not randomly assigned to participate in this study. All participants chose to participate, and thus, self-selection bias is possible.

Future research should include a control group and use random assignment to assign participants to either complete the cognitive training course or remain in a waitlist control group. This would ensure that self-selection bias is not present and would lead to confidence that changes observed with individuals who complete the cognitive training course would not have occurred in the absence of the course. Additionally, recruitment of future studies should involve obtaining a larger, more heterogeneous group of participants. This would ensure that effect sizes or inferential statistics are not unduly influenced by just one or two individuals and would ensure that findings can generalize to a larger proportion of the population.

The current study did not examine treatment adherence and competence or track attendance. Examining treatment adherence and competence involves examining whether the person administering the course is both competent in how to administer the course and whether they are following treatment protocols. All Mind Aerobics™ courses require that the person administering treatment to complete training and the program leader in the current study has successfully administered several courses before. Thus, I am confident that this individual is competent in how to administer the Lively Mind™ course. However, it would still be useful to officially measure adherence and competence. Tracking attendance would also have been a useful step to take in this study, as participants who do not attend the courses will not reap the benefits of the training.

Future research should address treatment adherence and competence. This would involve videotaping and having trained observers code videos of cognitive training sessions to determine whether the administration of the program adheres to the treatment manual. Additionally, quizzes would be administered to the person leading the Mind Aerobics™ course to ensure that they are competent in the material. This would lead to confidence that any effects observed are due to the intended model of the course.

A final limitation to note in this study is that two of the individuals in this study had co-morbid diagnoses; one had depression and one had depression and paranoia. These diagnoses were managed by medication however, and there was only one participant who scored a "6" on the PHQ-9 during the follow up test period (which is within the range of mild depressive symptoms). All other scores on the PHQ-9 were "3" or lower, which does not meet the threshold for mild depressive symptoms. Nevertheless, because depressive symptoms can come and go throughout the course of 6 months, it is possible that participants in this study experienced

depressive episodes on days outside of those during which assessment took place. Because depressive symptom often lead to difficulty in concentrating and can alter sleep schedules and overall well-being, depressive symptoms can certainly affect cognitive functioning.

An additional research question that should be addressed in the future is whether a comprehensive cognitive training program leads to the same level of cognitive gains as administering a cognitive training course that only focuses on a single domain. Because participants in the ACTIVE study (Ball et al., 2002) saw improvement in their respective cognitive domains, it is possible that completing an entire training course that only focuses on a single domain produces superior effects to a comprehensive course. The implication of this is that individuals with cognitive decline would want to identify particular problem domains and take courses focused on those domains.

References

Aarsland, D., Ballard, C., Walker, Z., Bostrom, F., Alves, G., Kossakowski, K., ... & Londos, E. (2009). Memantine in patients with Parkinson's disease dementia or dementia with Lewy bodies: a double-blind, placebo-controlled, multicentre trial. *The Lancet Neurology*, *8*(7), 613-618. doi:10.1186/isrctn89624516

Alzheimer's Association. (2017). 2017 Alzheimer's disease facts and figures report. *Alzheimer's Association*. Retrieved from http://www.alz.org/facts/.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Ashendorf, L., Jefferson, A. L., O'Connor, M. K., Chaisson, C., Green, R. C., & Stern, R. A. (2008). Trail Making Test errors in normal aging, mild cognitive impairment, and dementia. *Archives of Clinical Neuropsychology, 23*(2), 129-137. doi:10.1016/j.acn.2007.11.005

Bahar-Fuchs, A., Clare, L., & Woods, B. (2013). Cognitive training and cognitive rehabilitation for mild to moderate Alzheimer's disease and vascular dementia. *The Cochrane Library, 6*, 1-100. doi:10.1002/14651858.CD003260.pub2.

Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., ... & Unverzagt, F. W. (2002). Effects of cognitive training interventions with older adults. *JAMA, 288*(18), 2271-2281. doi:10.1001/jama.288.18.2271

Béland, F., Zunzunegui, M. V., Alvarado, B., Otero, A., & del Ser, T. (2005). Trajectories of cognitive decline and social relations. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*(6), 320-330. doi:10.1093/geronb/60.6.p320

Benedict, R. H. B., & Brandt, J. (2001). *Hopkins verbal learning test-revised (HVLT-R): Professional manual.* Lutz, FL: Psychological Assessment Resources.

Benedict, R. H. B., Schretlen, D., Groninger, L., Dobraski, M., & Shpritz, B. (1996). Revision of the brief visuospatial memory test: Studies of normal performance, reliability, and validity. *Psychological Assessment, 8*(2), 145-153. doi:http://dx.doi.org/10.1037/1040-3590.8.2.145

Boyle, L. L., Richardson, T. M., He, H., Xia, Y., Tu, X., Boustani, M., & Conwell, Y. (2011). How do the PHQ-2, the PHQ-9 perform in aging services clients with cognitive impairment? *International Journal of Geriatric Psychiatry*, *26*(9), 952-960. doi:10.1002/gps.2632

Clare, L., Linden, D. E., Woods, R. T., Whitaker, R., Evans, S. J., Parkinson, C. H., ... & Rugg, M. D. (2010). Goal-oriented cognitive rehabilitation for people with early-stage Alzheimer disease: a single-blind randomized controlled trial of clinical efficacy. *The American Journal of Geriatric Psychiatry*, *18*(10), 928-939. doi:10.1097/jgp.0b013e3181d5792a

Clare, L., & Woods, R. T. (2004). Cognitive training and cognitive rehabilitation for people with early-stage Alzheimer's disease: A review. *Neuropsychological Rehabilitation*, *14*(4), 385-401. doi:10.1080/09602010443000074

Coe, R. (2002). It's the effect size, stupid: What effect size is and why it is important. Retrieved from http://www.leeds.ac.uk/educol/documents/00002182.htm

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.) Hillsdale, NJ: Lawrence Earlbaum Associates.

Davis H. P., Klebe K. J., Guinther P. M., Schroder K. B., Cornwell R. E., James L. E. (2013). Subjective organization, verbal learning, and forgetting across the life span: From 5 to 89. *Experimental Aging Research 39*(1), 1–26. doi:10.1080/0361073X.2013.741956

Davis H. P., Small S. A., Stern Y., Mayeux R., Feldstein S. N., Keller F. R. (2003). Acquisition, recall, and forgetting of verbal information in long-term memory by young, middle-aged, and elderly individuals. *Cortex, 39*(4-5), 1063–1091. doi:10.1016/S0010-9452(08)70878-5

Emre, M., Tsolaki, M., Bonuccelli, U., Destée, A., Tolosa, E., Kutzelnigg, A., ... & 11018 Study Investigators. (2010). Memantine for patients with Parkinson's disease dementia or dementia with Lewy bodies: a randomised, double-blind, placebo-controlled trial. *The Lancet Neurology*, *9*(10), 969-977. doi:10.1016/S1474-4422(10)70194-0

Fox, C., Richardson, K., Maidment, I. D., Savva, G. M., Matthews, F. E., Smithard, D., ... & Brayne, C. (2011). Anticholinergic medication use and cognitive impairment in the older population: The medical research council cognitive function and ageing study. *Journal of the American Geriatrics Society*, *59*(8), 1477-1483. doi:10.1111/j.1532-5415.2011.03491.x

Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R. C., Ritchie, K., Broich, K., … Winblad, B. (2006). Mild cognitive impairment. *The Lancet, 367*(9518), 1262–1270. doi:10.1016/s0140-6736(06)68542-5

Giovagnoli, A. R., Del Pesce, M., Mascheroni, S., Simoncelli, M., Laiacona, M., & Capitani, E. (1996). Trail making test: normative values from 287 normal adult controls. *The Italian Journal of Neurological Sciences*, *17*(4), 305-309. doi:10.1007/bf01997792

Golden, C.J., & Freshwater, S.M. (2002). *Stroop color and word test- Adult version*. Wood Dale, IL: Stoelting.

Harada, C. N., Natelson Love, M. C., & Triebel, K. L. (2013). Normal cognitive aging. *Clinics in Geriatric Medicine, 29*(4), 737–752. doi:10.1016/j.cger.2013.07.002

Huckans, M., Hutson, L., Twamley, E., Jak, A., Kaye, J., & Storzbach, D. (2013). Efficacy of cognitive rehabilitation therapies for mild cognitive impairment (MCI) in older adults: Working toward a theoretical model and evidence-based interventions. *Neuropsychology Review, 23*(1), 63–80. doi:10.1007/s11065-013-9230-9

Inoue, M., Suhara, T., Sudo, Y., Okubo, Y., Yasuno, F., Kishimoto, T., ... & Tanada, S. (2001). Age-related reduction of extrastriatal dopamine D2 receptor measured by PET. *Life Sciences*, *69*(9), 1079-1084. doi:10.1016/S0024-3205(01)01205-X

Jak, A. J., Bondi, M. W., Delano-Wood, L., Wierenga, C., Corey-Bloom, J., Salmon, D. P., & Delis, D. C. (2009). Quantification of five neuropsychological approaches to defining mild cognitive impairment. *The American Journal of Geriatric Psychiatry, 17*(5), 368–375. doi:10.1097/jgp.0b013e31819431d5

Kanaan, S. F., McDowd, J. M., Colgrove, Y., Burns, J. M., Gajewski, B., & Pohl, P. S. (2014). Feasibility and efficacy of intensive cognitive training in early-stage Alzheimer's Disease. *American Journal of Alzheimer's Disease & Other Dementias*, 29(2), 150–158. doi:10.1177/1533317513506775

Koss, E., Ober, B. A., Delis, D. C., & Friedland, R. P. (1984). The Stroop color-word test: Indicator of dementia severity. *International Journal of Neuroscience*, *24*(1), 53-61. doi:10.3109/00207458409079534

Krishna R., Moustafa A. A., Eby L. A., Skeen L. C., Myers C. E. (2012). Learning and generalization in healthy aging: implication for frontostriatal and hippocampal function. *Cognitive and Behavioral Neurology, 25*(1), 7–15. doi:10.1097/WNN.0b013e318248ff1b

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals, 32*(9), 509-515. doi:10.3928/0048-5713-20020901-06

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606-613. doi:10.1046/j.1525-1497.2001.016009606

Mate-Kole, C. C., Fellows, R. P., Said, P. C., McDougal, J., Catayong, K., Dang, V., & Gianesini, J. (2007). Use of computer assisted and interactive cognitive training programmes with moderate to severely demented individuals: a preliminary study. *Aging & Mental Health,11*(5), 485-495. doi:10.1080/13607860701368422

McDowell, I., Kristjansson, B., Hill, G. B., & Hébert, R. (1997). Community screening for dementia: The mini mental state exam (MMSE) and modified mini-mental state exam (3MS) compared. *Journal of Clinical Epidemiology*, *50*(4), 377-383. doi:10.1016/s0895-4356(97)00060-7

Mild cognitive impairment. (n.d.). *Mayo Clinic*. Retrieved from https://www.mayoclinic.org/diseases-conditions/mild-cognitive-impairment/symptoms-causes/syc-20354578

Mordkoff, T. (2000). The assumption of normality. Retrieved from http://www2.psychology.uiowa.edu/faculty/mordkoff/GradStats/part%201/I.07%20normal.pdf

Mori, E., Ikeda, M., & Kosaka, K. (2012). Donepezil for dementia with Lewy bodies: A randomized, placebo-controlled trial. *Annals of neurology*, *72*(1), 41-52. doi:10.1002/ana.23557

Morris, M. C. (2012). Nutritional determinants of cognitive aging and dementia. *Proceedings of the Nutrition Society*, *71*(1), 1-13. doi:10.1017/s0029665111003296

Moustafa, A. A., Hewedi, D. H., Eissa, A. M., Myers, C. E., & Sadek, H. A. (2012). The relationship between associative learning, transfer generalization, and homocysteine levels in mild cognitive impairment. *PLoS One*, *7*(9), e46496. doi:10.1371/journal.pone.0046496

New England Cognitive Center. (2015). *Lively mind program sourcebook.* (2nd ed.). New London, CT: New England Cognitive Center, Inc.

Nordahl, C. W., Ranganath, C., Yonelinas, A. P., DeCarli, C., Fletcher, E., & Jagust, W. J. (2006). White matter changes compromise prefrontal cortex function in healthy elderly individuals. *Journal of Cognitive Neuroscience*, *18*(3), 418-429. doi:10.1162/089892906775990552

Rapp, S., Brenes, G., & Marsh, A. P. (2002). Memory enhancement training for older adults with mild cognitive impairment: A preliminary study. *Aging & Mental Health, 6*(1), 5–11. doi:10.1080/13607860120101077

Rebok, G. W., Ball, K., Guey, L. T., Jones, R. N., Kim, H. Y., King, J. W., ... & Willis, S. L. (2014). Ten-year effects of the advanced cognitive training for independent and vital elderly cognitive training trial on cognition and everyday functioning in older adults. *Journal of the American Geriatrics Society*, *62*(1), 16-24. doi:10.1111/jgs.12607

Reitan, R. M. (1959). Correlations between the trail making test and the Wechsler-Bellevue scale. *Perceptual and Motor Skills*, *9*(3), 127-130. doi:10.2466/pms.9.3.127-130

Ross, T. P., Calhoun, E., Cox, T., Wenner, C., Kono, W., & Pleasant, M. (2007). The reliability and validity of qualitative scores for the controlled oral word association test. *Archives of Clinical Neuropsychology*, *22*(4), 475-488. doi:10.1016/j.acn.2007.01.026

Schretlen, D., Bobholz, J. H., & Brandt, J. (1996). Development and psychometric properties of the Brief Test of Attention. *The Clinical Neuropsychologist*, *10*(1), 80-89. doi:10.1080/13854049608406666

Schneider, L. S., Dagerman, K. S., Higgins, J. P., & McShane, R. (2011). Lack of evidence for the efficacy of memantine in mild Alzheimer disease. *Archives of Neurology*, *68*(8), 991-998. doi:10.1001/archneurol.2011.69

Shapiro, A. M., Benedict, R. H., Schretlen, D., & Brandt, J. (1999). Construct and concurrent validity of the Hopkins verbal learning test–revised. *The Clinical Neuropsychologist*, *13*(3), 348-358. doi:10.1076/clin.13.3.348.1749

Small S. A., Schobel S. A., Buxton R. B., Witter M. P., Barnes C. A. (2011). A pathophysiological framework of hippocampal dysfunction in ageing and disease. *Nature Reviews Neuroscience 12*(10), 585–601. doi:10.1038/nrn3085

Snitz, B. E., O'Meara, E. S., Carlson, M. C., Arnold, A. M., Ives, D. G., Rapp, S. R., ... & DeKosky, S. T. (2009). Ginkgo biloba for preventing cognitive decline in older adults: A randomized trial. *JAMA*, *302*(24), 2663-2670. doi:10.1001/jama.2009.1913

Stypulkowski, K. (2017). *Addressing cognitive decline: Evaluating the effects of a cognitive training program for individuals with mild to moderate cognitive impairment.* (Unpublished Master's Thesis). Minnesota State University, Mankato, Mankato, MN.

Tombaugh, T. N., McDowell, I., Kristjansson, B., & Hubley, A. M. (1996). Mini-mental state examination (MMSE) and the modified MMSE (3MS): A psychometric comparison and normative data. *Psychological Assessment*, *8*(1), 48-59. doi:10.1037/1040-3590.8.1.48

Trinh, N. H., Hoblyn, J., Mohanty, S., & Yaffe, K. (2003). Efficacy of cholinesterase inhibitors in the treatment of neuropsychiatric symptoms and functional impairment in Alzheimer disease: A meta-analysis. *JAMA*, *289*(2), 210-216. doi:10.1001/jama.289.2.210

Wechsler, D. (2008). *Wechsler adult intelligence scale–fourth edition* (WAIS–IV). San Antonio, TX: NCS Pearson.

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, 296*(23), 2805-2814. doi:10.1001/jama.296.23.2805

Woods, B., Aguirre, E., Spector, A. E., & Orrell, M. (2012). Cognitive stimulation to improve cognitive functioning in people with dementia. *The Cochrane Library*. doi:10.1002/14651858.cd005562.pub2

United Nations. (2017). *Probabilistic projections based on the World Population Prospects: The 2017 Revision*. [Excel file]. Retrieved from https://esa.un.org/unpd/wpp/Download/Probabilistic/Population/

United States Census Bureau. (2017). *Facts for features: Older Americans month: May 2017* [Fact sheet]. Retrieved from https://www.census.gov/newsroom/facts-for-features/2017/cb17-ff08.html

Yaffe, K., Barnes, D., Nevitt, M., Lui, L.-Y., & Covinsky, K. (2001). A prospective study of physical activity and cognitive decline in elderly women. *Archives of Internal Medicine, 161*(14), 1703. doi:10.1001/archinte.161.14.1703

Zelinski, E. M., & Gilewski, M. J. (2004). A 10-item Rasch modeled memory self-efficacy scale. *Aging & Mental Health*, *8*(4), 293-306. doi:10.1080/13607860410001709665

Table 1.
Pre-test means for those with and without 3MS scores between 60-80.

Assessment	Pre-test Mean (SD) 3MS 60-80	Pre-test Mean (SD) 3MS above 80	Effect Size (Cohen's d)	Confidence Interval	Interpretation of Effect Size
3MS	72.33 (2.89)	89.5 (2.12)	6.48	2.08 - 10.87	Large Effect
HVLT-R Total Recall	10.33 (2.08)	17.50 (0.71)	4.10	0.99 - 7.21	Large Effect
HVLT-R Delayed Recall	0.33 (0.58)	7.00 (2.83)	3.92	0.90 - 6.94	Large Effect
HVLT-R Recognition	4.00 (1.73)	7.00 (2.83)	1.39	-0.60 - 3.38	Large Effect
Forward Digit Span Correct	8.00 (1.00)	8.00 (1.41)	0.00	-1.79 - 1.79	No Effect
Backward Digit Span Correct	7.67 (2.89)	7.50 (0.71)	-0.07	-1.86 - 1.72	No Effect
BTA Total	10.33 (2.31)	9.50 (0.71)	-0.43	-2.24 - 1.38	Small Effect
Trails A	43.67 (23.25)	44.00 (8.48)	0.02	-1.77 - 1.81	No Effect
Stroop Test Word	87.33 (12.58)	56.50 (27.58)	-1.63	-3.68 - 0.43	Large Effect
Stroop Test Color	44.00 (15.59)	44.50 (0.71)	0.04	-1.75 - 1.83	No Effect
Stroop Test Color-Word	15.67 (2.31)	18.00 (4.24)	0.75	-1.10 - 2.60	Intermediate Effect
COWAT Total	29.33 (5.51)	18.50 (17.68)	-0.97	-2.86 - 0.92	Large Effect
BVMT-R Total Recall	6.33 (1.53)	16.00 (4.24)	3.52	0.69 - 6.34	Large Effect
BVMT-R Delayed Recall	2.00 (1.00)	5.00 (4.24)	1.16	-0.77 - 3.09	Large Effect
BVMT-R Recognition	2.67 (1.53)	5.00 (1.41)	1.56	-0.47 - 3.60	Large Effect
Visual Puzzles Total	10.00 (3.46)	10.00 (1.41)	0.00	-1.79 - 1.79	No Effect
Frequency of Forgetting Rasch Scale	0.12 (0.43)	0.82 (1.01)	1.03	-0.87 - 2.93	Large Effect
PHQ-9	1.33 (1.53)	1.00 (1.41)	-0.22	-2.02 - 1.57	Small Effect

Table 2.
Means for all participants on pre-test and post-test (N=5).

Assessment	Pre-test Mean (SD)	Post-test Mean (SD)	Effect Size (Cohen's d)	Confidence Interval	Interpretation of Effect Size
3MS	79.20 (9.68)	84.80 (6.76)	0.99	-0.32 - 2.31	Large Effect
HVLT-R Total Recall	13.20 (4.20)	13.40 (3.78)	0.09	-1.15 - 1.33	No Effect
HVLT-R Delayed Recall	3.00 (3.94)	2.60 (2.79)	-0.34	-1.59 - 0.91	Small Negative Effect
HVLT-R Recognition	5.20 (2.49)	4.80 (2.39)	-0.16	-1.40 - 1.08	No Effect
Forward Digit Span Correct	8.00 (1.00)	7.80 (0.84)	-0.44	-1.69 - 0.82	Small Negative Effect
Backward Digit Span Correct	7.60 (2.07)	8.20 (2.17)	0.41	-0.85 - 1.66	Small Effect
BTA Total	10.00 (1.73)	8.00 (2.91)	-8.18	-11.97 - -4.38	Large Negative Effect
Trails A	43.80 (16.98)	58.00 (28.12)	1.69	0.24 - 3.13	Large Effect (Decline)
Stroop Test Word	75.00 (23.54)	78.80 (14.86)	0.13	-1.12 - 1.37	No Effect
Stroop Test Color	44.20 (11.03)	43.40 (12.62)	-0.08	-1.32 - 1.16	No Effect
Stroop Test Color-Word	16.60 (2.97)	11.60 (5.08)	-2.78	-4.51 - -1.04	Large Negative Effect
COWAT Total	25.00 (11.34)	15.80 (6.91)	-4.54	-6.88 - -2.19	Large Negative Effect
BVMT-R Total Recall	10.20 (5.81)	12.20 (8.07)	1.07	-0.26 - 2.39	Large Effect
BVMT-R Delayed Recall	3.20 (2.77)	3.00 (3.16)	-0.15	-1.39 - 1.09	No Effect
BVMT-R Recognition	3.60 (1.82)	3.40 (1.14)	-0.10	-1.34 - 1.15	No Effect
Visual Puzzles Total	10.00 (2.55)	10.20 (2.77)	0.20	-1.05 - 1.44	Small Effect
Frequency of Forgetting Rasch Scale	0.40 (0.70)	0.19 (0.32)	-0.38	-1.63 - 0.87	Small Negative Effect
PHQ-9	1.20 (1.30)	1.80 (1.30)	0.40	-0.86 - 1.65	Small Effect

Table 3.
Means for all participants on post-test and follow-up (N=5).

Assessment	Post-test Mean (SD)	Follow-up Mean (SD)	Effect Size (Cohen's d)	Confidence Interval	Interpretation of Effect Size
3MS	84.80 (6.76)	80.80 (7.29)	-1.08	-2.41 - 0.25	Large Negative Effect
HVLT-R Total Recall	13.40 (3.78)	14.20 (3.56)	0.23	-1.01 - 1.48	Small Effect
HVLT-R Delayed Recall	2.60 (2.79)	1.80 (2.68)	-0.72	-2.00 - 0.56	Intermediate Negative Effect
HVLT-R Recognition	4.80 (2.39)	5.20 (2.77)	0.40	-0.85 - 1.65	Small Effect
Forward Digit Span Correct	7.80 (0.84)	8.40 (1.81)	0.76	-0.52 - 2.04	Intermediate Effect
Backward Digit Span Correct	8.20 (2.17)	8.20 (0.84)	0.00	-1.24 - 1.24	No Effect
BTA Total	8.00 (2.91)	9.40 (6.07)	1.48	0.08 - 2.88	Large Effect
Trails A	58.00 (28.12)	47.20 (19.30)	-0.86	-2.16 - 0.43	Large Negative Effect (Improvement)
Stroop Test Word	78.80 (14.86)	75.20 (19.03)	-0.35	-1.59 - 0.90	Small Negative Effect
Stroop Test Color	43.40 (12.62)	37.80 (16.44)	-1.21	-2.56 - 0.14	Large Negative Effect
Stroop Test Color-Word	11.60 (5.08)	14.00 (5.39)	0.47	-0.78 - 1.73	Small Effect
COWAT Total	15.80 (6.91)	15.80 (6.83)	0.00	-1.24 - 1.24	No Effect
BVMT-R Total Recall	12.20 (8.07)	13.20 (10.52)	0.53	-0.73 - 1.80	Intermediate Effect
BVMT-R Delayed Recall	3.00 (3.16)	2.20 (3.90)	-0.39	-1.64 - 0.86	Small Negative Effect
BVMT-R Recognition	3.40 (1.14)	4.20 (1.10)	0.94	-0.37 - 2.25	Large Effect
Visual Puzzles Total	10.20 (2.77)	12.40 (4.56)	1.41	0.03 - 2.80	Large Effect
Frequency of Forgetting Rasch Scale	0.19 (0.32)	0.42 (0.55)	0.60	-0.66 - 1.87	Intermediate Effect
PHQ-9	1.80 (1.30)	2.00 (2.55)	0.11	-1.14 - 1.35	No Effect

Table 4.
Means for participants with 3MS scores between 60-80 on pre-test and post-test (n=3).

Assessment	Pre-test Mean (SD)	Post-test Mean (SD)	Effect Size (Cohen's d)	Confidence Interval	Interpretation of Effect Size
3MS	72.33 (2.89)	80.33 (3.79)	1.67	-0.19 - 3.52	Large Effect
HVLT-R Total Recall	10.33 (2.08)	11.00 (1.73)	0.27	-1.34 - 1.88	Small Effect
HVLT-R Delayed Recall	0.33 (0.58)	0.67 (1.15)	*Not Available	*Not Available	*Not Available
HVLT-R Recognition	4.00 (1.73)	3.33 (1.53)	-0.20	-1.80 - 1.41	Small Negative Effect
Forward Digit Span Correct	8.00 (1.00)	7.67 (0.58)	-0.64	-2.28 - 1.00	Intermediate Negative Effect
Backward Digit Span Correct	7.67 (2.89)	7.67 (2.08)	0.00	-1.60 - 1.60	No Effect
BTA Total	10.33 (2.31)	8.67 (3.79)	-5.36	-8.78 - -1.93	Large Negative Effect
Trails A	43.67 (23.25)	54.67 (31.79)	1.41	-0.38 - 3.20	Large Effect (Decline)
Stroop Test Word	87.33 (12.58)	83.00 (14.80)	-0.68	-2.33 - 0.96	Intermediate Negative Effect
Stroop Test Color	44.00 (15.59)	45.33 (14.64)	0.13	-1.47 - 1.73	No Effect
Stroop Test Color-Word	15.67 (2.31)	11.33 (5.13)	-2.56	-4.72 - -0.40	Large Negative Effect
COWAT Total	29.33 (5.51)	18.00 (4.58)	-8.54	-13.63 - -3.45	Large Negative Effect
BVMT-R Total Recall	6.33 (1.53)	6.33 (0.58)	0.00	-1.6 - 1.6	No Effect
BVMT-R Delayed Recall	2.00 (1.00)	1.00 (1.00)	-1.00	-2.70 - 0.70	Large Negative Effect
BVMT-R Recognition	2.67 (1.53)	2.67 (0.58)	0.00	-1.6 - 1.6	No Effect
Visual Puzzles Total	10.00 (3.46)	11.00 (3.46)	*Not Available	*Not Available	*Not Available
Frequency of Forgetting Rasch Scale	0.12 (0.43)	0.06 (0.20)	-0.08	-1.68 - 1.52	No Effect
PHQ-9	1.33 (1.53)	1.67 (1.53)	0.59	-1.05 - 2.22	Intermediate Effect (Decline)

Table 5.
Means for participants with 3MS scores between 60-80 on post-test and follow-up (n=3).

Assessment	Post-test Mean (SD)	Follow-up Mean (SD)	Effect Size (Cohen's d)	Confidence Interval	Interpretation of Effect Size
3MS	80.33 (3.79)	76.66 (6.42)	-0.99	-2.68 - 0.71	Large Negative Effect
HVLT-R Total Recall	11.00 (1.73)	13.00 (4.36)	1.05	-0.66 - 2.76	Large Effect
HVLT-R Delayed Recall	0.67 (1.15)	0.00 (0.00)	-0.41	-2.03 - 1.21	Small Negative Effect
HVLT-R Recognition	3.33 (1.53)	4.00 (3.00)	2.31	0.24 - 4.37	Large Effect
Forward Digit Span Correct	7.67 (0.58)	9.00 (1.73)	2.29	0.23 - 4.35	Large Effect
Backward Digit Span Correct	7.67 (2.08)	8.33 (0.58)	1.32	-0.45 - 3.08	Large Effect
BTA Total	8.67 (3.79)	11.67 (7.37)	10.22	4.22 - 16.22	Large Effect
Trails A	54.67 (31.79)	44.33 (25.93)	-1.18	-2.91 - 0.55	Large Negative Effect
Stroop Test Word	83.00 (14.80)	82.00 (23.39)	-0.53	-2.16 - 1.09	Intermediate Negative Effect
Stroop Test Color	45.33 (14.64)	41.00 (21.28)	-0.99	-2.68 - 0.71	Large Negative Effect
Stroop Test Color-Word	11.33 (5.13)	12.67 (6.51)	0.23	-1.38 - 1.83	Small Effect
COWAT Total	18.00 (4.58)	19.67 (3.51)	3.33	0.86 - 5.8	Large Effect
BVMT-R Total Recall	6.33 (0.58)	5.67 (2.52)	-1.81	-3.71 - 0.09	Large Negative Effect
BVMT-R Delayed Recall	1.00 (1.00)	0.67 (1.15)	-0.17	-1.77 - 1.43	No Effect
BVMT-R Recognition	2.67 (0.58)	3.67 (1.15)	1.72	-0.15 - 3.60	Large Effect
Visual Puzzles Total	11.00 (3.46)	14.33 (5.03)	1.53	-0.29 - 3.35	Large Effect
Frequency of Forgetting Rasch Scale	0.06 (0.20)	0.19 (0.58)	0.35	-1.27 - 1.96	Small Effect
PHQ-9	1.67 (1.53)	1.33 (1.53)	-0.59	-2.22 - 1.05	Intermediate Negative Effect (Improvement)

Table 6.
Raw scores for all participants on measures that showed large effect sizes.

Participant Number	3MS Pre-test Score (Qualifying Status)	BTA	COWAT	Stroop Color-Word
ZLX001	69 (Qualifies)	Pre- 9 Post- 6 Follow Up- 6	Pre- 33 Post-22 Follow Up- 23	Pre- 13 Post-7 Follow Up-6
WJX002	74 (Qualifies)	Pre- 9 Post- 7 Follow Up- 9	Pre- 23 Post-13 Follow Up- 16	Pre- 17 Post-17 Follow Up- 13
RJX004	74 (Qualifies)	Pre- 13 Post-13 Follow Up- 20	Pre- 32 Post-19 Follow Up- 20	Pre- 17 Post-10 Follow Up- 19
PBX005	88 (Does not qualify)	Pre- 10 Post- 8 Follow Up- 6	Pre- 31 Post-20 Follow Up- 15	Pre- 21 Post-17 Follow Up- 19
MMX003	91 (Does not qualify)	Pre- 9 Post- 6 Follow Up- 6	Pre- 6 Post-5 Follow Up- 5	Pre- 15 Post-7 Follow Up- 13