FEASIBILITY AND EFFECTIVENESS OF A TELEHEALTH-DELIVERED INDUCTIVE REASONING TRAINING PROGRAM FOR OLDER ADULTS

By

BRAD P. TAYLOR

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2023

To my family, friends, and mentors, all who have supported and inspired my career interests in cognitive aging

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (NIA T32 AG020499). Research reported in this publication was supported by the University of Florida Clinical and Translational Science Institute, which is supported in part by the NIH National Center for Advancing Translational Sciences under award number UL1TR001427. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Access to data from the ACTIVE study is currently supported by NIA R01 AG056486. The ACTIVE Cognitive Training Trial was supported by grants from the National Institutes of Health to six field sites and the coordinating center, including: Hebrew Senior-Life, Boston (NR04507), Indiana University School of Medicine (NR04508), Johns Hopkins University (AG014260), New England Research Institutes (AG014282), Pennsylvania State University (AG14263), University of Alabama at Birmingham (AG14289), and University of Florida (AG014276).

I wish to thank my chair and mentor, Dr. Michael Marsiske, who brought his guidance, encouragement, and support at every step of my UF experience. I also wish to thank Jeff Boissoneault, Ph.D., Carol Mathews, M.D., and Glenn Smith, Ph.D., ABPP-CN, who provided invaluable advice as members of my dissertation committee. I am grateful to my Quantitative Aging and Neurocognitive Training (QuANT) Labmates (Jacqueline Maye, Lindsay Rotblatt, Joshua Owens, and Jacob Fiala) and my graduate cohort, all of whom I consider my lifelong friends and colleagues. I also want to express my sincere thanks to my nine wonderful Research Assistants (Madison Verdone, Ryan Faulkner, Barbara Balsamo, Christy Silver, Lane Fontana, Lauren Cox, Rachel Hamaker, Brianna Nazario, and Abigail Chalfant). A special thank you goes to the older adults who participated in my study. Their warmth and kindness helped make this an enjoyable study and learning experience. I also wish to thank Lana Ying for her support in the propensity analysis. A final thanks to Binh Nguyen, who was my collaborator and provided productive discussions during an earlier phase of this project.

My family, particularly Denise Taylor, Ralph Taylor, Esma Taylor, and Katie Cox have all provided me the needed encouragement to help me accomplish my dissertation and my PhD.

TABLE OF CONTENTS

APPENDIX

LIST OF TABLES

Table page

LIST OF FIGURES

Figure page

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

FEASIBILITY AND EFFECTIVENESS OF A TELEHEALTH-DELIVERED INDUCTIVE REASONING TRAINING PROGRAM FOR OLDER ADULTS

By

Brad Taylor

August 2023

Chair: Michael Marsiske
Major: Psychology

Older age is a period of risk for decline in inductive reasoning, an indicator of fluid intelligence. Inductive reasoning training has been found to be particularly effective at improving inductive reasoning, with some evidence of improved everyday functioning and driving. Additionally, telehealth interventions have been found effective for mood disorders. The present study investigated the feasibility and effectiveness of an inductive reasoning training program, designed to mimic the inductive reasoning arm of the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) trial, over a telehealth-delivered format (Zoom and Canvas). 31 older adult participants received 10-sessions of telehealth-delivered inductive reasoning training over 5-weeks. Comparison groups (inductive reasoning trained and no-contact controls) were culled from the in-person ACTIVE trial via propensity matching. All participants completed pretest and posttest inductive reasoning tasks and questionnaires. Participants in the telehealth training program rated the program highly and demonstrated similar inductive

reasoning gains from pretest to posttest as the matched participants from the original, in-person inductive reasoning training in ACTIVE on 2 of the 3 inductive reasoning measures. Word Series performance gains were less in the telehealth trained group when compared to ACTIVE’s in-person trained participants and were similar to those of matched no-contact controls from ACTIVE. This discrepancy may be due to telehealth limitations. Broadly speaking, much of the improvement in inductive reasoning performance in both training groups appears to reflect increases in the number of attempted problems, suggesting similar mechanisms of improvement (accuracy and efficiency) between the two training groups. Overall, this study suggests that telehealth delivery may be a viable form of cognitive training in inductive reasoning.

CHAPTER 1
STATEMENT OF THE PROBLEM

Inductive reasoning, or the ability to “educe novel relations”, is an important indicator of fluid intelligence and has long been shown to evince negative effects of age in cross-sectional (Horn & Cattell, 1967) and longitudinal (Schaie, 1996) studies. More recently, inductive reasoning has emerged as an important predictor of older adults’ everyday problem solving (Gross et al., 2011; Willis et al., 1992; Yam et al., 2014).

One reason for an enduring focus on inductive reasoning over the past five decades is that inductive reasoning demonstrates substantial and durable benefits of 5-10 sessions of cognitive training in older adults (Kelly et al., 2014). Benefits have primarily been seen directly on measures of inductive reasoning (Boron et al., 2007; Willis & Caskie, 2013; Willis & Schaie, 1986), but there is some evidence from Rebok and colleagues of transfer to perceived everyday functioning (2014) and driving (Ball et al., 2010). A challenge in scaling inductive reasoning training broadly, however, is that existing training studies have relied on in-person intervention (Ball et al., 2002; Jobe et al., 2001); this is both labor intensive and requires trainers with specialized training and facilities in which to conduct training. With expanded awareness and use of telehealth for psychological intervention delivery, this study sought to examine whether telehealth-delivered inductive reasoning training is feasible and effective. There is evidence that telehealth can be an effective approach to delivering psychological and behavioral interventions (Burton & O'Connell, 2018; Holmqvist et al., 2014; Khatri et al., 2014), including cognitive rehabilitation and compensatory memory skills training (Burton & O'Connell, 2018; Lawson et al., 2020); however, this cognitive training intervention in inductive reasoning had not yet been adapted for telehealth delivery.

Older adults as a group may find the receipt of psychological interventions via telehealth to be challenging. Nevertheless, telehealth cognitive training interventions also have the unique opportunity to increase ecological validity by being done in the participants' homes (Margrett & Willis, 2006), to increase access for older adults (Demiris et al., 2009) by reducing transportation and physical mobility requirements, and to keep participants and trainers safe during the COVID-19 outbreak (Monaghesh & Hajizadeh, 2020) or other future public health emergencies. Accordingly, in an older sample, this study assessed the effect of telehealth-delivered inductive reasoning training on participant perceptions of usefulness, ease of use & learnability, interface quality, reliability, satisfaction and intended future use. The study assessed whether older participants evinced improvement in inductive reasoning following telehealth-delivered cognitive training in this domain. Each older adult participant received 5 weeks of training over ten sessions, with pretest and posttest batteries administered before and after training. Telehealth-delivered inductive reasoning training was adapted from the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) study's in-person inductive reasoning training program, and focused on improving the ability to solve problems that require reasoning about patterns that follow a serial pattern or sequence (Jobe et al., 2001).

The study also leveraged a large available data set to permit comparison of telehealth training effects to several other conditions. The study used two propensity-matched comparison groups from the ACTIVE study, one was drawn from 699 participants who received face-to-face in-person group inductive reasoning training, and the second was drawn from 698 participants assigned to a no-contact control group.

The chief contribution of the current study was to explore the feasibility and preliminary relative effectiveness of a telehealth-delivered inductive reasoning intervention. The results of the study help to provide insight on whether a telehealth-adapted reasoning training shows promise with older adults and can be used to power future clinical trial studies to more definitively evaluate the effectiveness of telehealth delivered training.

The present study had two specific aims:

Aim 1: To determine client perceptions of and compliance with a telehealth-delivered cognitive training intervention in reasoning.

Hypothesis 1.1 It was hypothesized that subjects would provide “agree” or better ratings on questions asking for their perceptions of the usefulness, ease of use & learnability, interface quality, interaction quality, reliability, and satisfaction & future use of the telehealth reasoning training program.

Hypothesis 1.2 It was hypothesized that participants would attend at least 80% (8/10) of training sessions, on average, a rate comparable to that observed in in-person trials.

Aim 2: To determine if a telehealth cognitive training intervention in inductive reasoning was effective in improving inductive reasoning.

Hypothesis 2.1 It was hypothesized that those who received inductive reasoning telehealth training would perform significantly better at an immediate post-test that occurred about one week after the conclusion of a 5-week, 10-session training program.

Hypothesis 2.2 Participants' training gain would exceed simple practice effects, as evidenced by telehealth-trained individuals showing greater gain than a matched no-contact control group culled from the ACTIVE clinical trial.

Aim 3: To conduct a descriptive comparison to examine whether a telehealth cognitive training intervention in inductive reasoning was comparable to traditional face-to-face inductive reasoning training.

Hypothesis 3.1 Participants receiving telehealth training would experience training gains similar to those observed in participants who received traditional face-to-face inductive reasoning training. These results would provide us with effect size estimates to power a future equivalency trial.

CHAPTER 2
LITERATURE REVIEW

Overview

The following sections provide 1) a review of cognitive training, including consideration of transfer, durability, and dementia prevention, as well as a theoretical consideration of restoration versus compensation as the goals of training; 2) a specific review of inductive reasoning training; 3) a consideration of the ACTIVE study, whose sample is the source of comparison groups for the study; and 4) an exploration of telehealth with older adults and how it may help to circumvent some of the limitations of in-person cognitive training paradigms. Chapter 2 concludes with a summary of research best practices for cognitive training paradigms and how our intervention study adhered to these best practice recommendations.

At the highest level, the motivation for cognitive training studies with older adults comes from the observation that the global]population is increasing in age (Chiu et al., 2017; Kinsella & Phillips, 2005), and the growth of older cohorts is associated with a concomitant growth in the prevalence of dementia. Normative aging is associated with widespread decline in many areas of cognition (Craik & Salthouse, 2011), and this is a source of concern for many elders. Nearly half of community-dwelling older adults report cognitive concerns (Dark-Freudeman et al., 2006; Lowenthal & Berkman, 1967). Considering the continued aging of the population, the prevalence of dementia, an age-related disorder, has also continued to increase (Smith & Bondi, 2013; Unverzagt et al., 2012). In fact, the prevalence of Alzheimer's disease, by far the leading cause of dementia (Morris, 1996), is expected to increase to ~8-13 million persons by the year 2050, representing a fourfold increase from the prevalence in 2000 (Sloane et al., 2002). Dementia affects not only the individual with the condition, but also their families and the economy, with the global cost of dementia believed to be approximately $1 trillion USD annually (Livingston et al., 2020; Patterson, 2018). Therefore, identifying potential interventions to delay the onset of dementia and maintain cognitive functioning and functional independence are of enormous public health significance (Unverzagt et al., 2012).

Cognitive Aging

Cognitive aging is a lifespan process involving gradual, yet multidimensional and multidirectional changes in cognitive functioning that occur as people age (Baltes, 1987; Blazer et al., 2015). These multifaceted changes encompass both growth and decline (Baltes, 1997), with fluid abilities that rely on rapid encoding and modification of new information often showing steeper declines across adulthood (and especially in the later decades), while crystallized abilities that rely on culture/knowledge may remain stable or even improve into older age (Stine-Morrow, 2007). Some of the crystallized domains of intelligence include domains of vocabulary, general information, and reading comprehension. The fluid, age-sensitive domains include episodic memory, general visualization/spatial orientation, attention, processing speed, executive functioning, and reasoning. However, cognitive aging is not a monolithic process that affects everyone equally, as there are often large individual differences in lifespan cognitive trajectories (Lindenberger & von Oertzen, 2006).

Cognitive training as a type of behavioral intervention

Once we acknowledge that there may be individual differences in cognitive aging, it follows that interventions might be one way of altering individual trajectories; indeed, this idea of cognitive interventions as trajectory-modifying strategies was first offered by Nancy Denney (Denney, 1984; Hertzog et al., 2008) . Thus, behavioral interventions may be a means of modifying cognitive aging trajectories (Dixon & Lachman, 2018; Raz, 2009), thereby helping to preserve cognition and maintain functional independence (Smith, 2016). Under the umbrella of behavioral interventions exists cognitive-based training, which is designed to support the maintenance of cognitive functions (Chiu et al., 2017; Winocur et al., 2007). While no consensus lexicon yet describes the various ways in which cognition can be modified to reach desired aims, much of the literature on intellectual activities designed to maintain cognitive functioning can be divided into three categories: cognitive stimulation, cognitive rehabilitation, and cognitive training (Steinerman, 2010).

Cognitive training, which, as its name implies, often involves training on one or more cognitive tasks (Simons et al., 2016), refers to theory-guided intervention that often involves direct, purposive instruction of some form of cognitive strategy (Smith, 2016) on one or more targeted cognitive abilities. In addition to strategic instruction, cognitive training also frequently employs practice exercises, feedback, adaptive difficulty, and small groups for peer support. The use of these cognitive strategies is believed to be one of the main mechanisms by which cognitive change can occur (Willis & Schaie, 2009).

Can older adults profit from cognitive training? And in what domains?

Systematic cognitive training in older adults has been shown to produce long-term improvements in cognitive functioning (Ball et al., 2002; Sherry L. Willis et al., 2006); however, evidence regarding the prevention or delay of dementia is insufficient (Butler et al., 2018). Several systematic reviews and meta-analyses have indicated that cognitive training in healthy older adults has a moderate effect on overall cognitive function (Chiu et al., 2017), such that training improves cognitive performance on the trained domain. Evidence regarding whether it can prevent/delay cognitive decline or dementia outcomes are presently insufficient, largely due to the absence of high quality, long-term trials (Butler et al., 2018). Correspondingly, a systematic review team from the World Health Organization (WHO), which assessed the quality of existing reviews, determined that the quality of evidence supporting cognitive training for reducing dementia risk in healthy older adults is “low” (2019).

Another interdisciplinary, international expert consensus panel also evaluated the quality of evidence regarding cognitive training (Livingston et al., 2020). Basing their conclusion upon the evidence provided in three recent systematic reviews (Butler et al., 2018; Gates et al., 2019; Kane et al., 2017), they concluded that cognitive training has the ability to improve the trained domain (Livingston et al., 2020), with unknown benefits for function and dementia. In two reviews specifically examining the effects of computerized cognitive training in healthy older adults, similar findings were noted, such that computerized cognitive training was modestly effective for improving their cognitive performance (Kueider et al., 2012; Lampit et al., 2014).

Looking at the literature on cognitive training in those with mild cognitive impairment (MCI), a systematic review and meta-analysis investigating the everyday impact of cognitive intervention in mild cognitive impairment (MCI) found small median effects for ADLs, mood, and metacognitive outcomes (Chandler et al., 2016). By comparison, another meta-analysis of the efficacy of cognitive interventions in those with MCI on neuropsychological outcomes concluded that individuals with MCI who received multicomponent training on multiple domains demonstrated improved cognitive outcomes post-intervention (Sherman et al., 2017). Looking again at computerized cognitive training, but in older adults with MCI or dementia, a review found that computerized cognitive training was effective at improving select cognitive domains in individuals with MCI, but evidence for efficacy in those with dementia was weak (Hill et al., 2017).

Transfer defined, and evidence of cognitive training transfer

Societally, there is a shared belief that for "cognitive interventions" in early life (i.e., formal education) to have durable effects and to transfer to everyday life, it must be extended in time, adaptively increasing in difficulty and complexity, with teaching aimed at explicitly helping trainees to generalize to everyday life. Thus, the value of schooling relies, in part, on the belief that learning and the application of specific learned content may generalize to non-academic contexts (Simons et al., 2016). Yet, even in formal education, there is often evidence of limited transfer or generalization to real-world tasks (Billing, 2007; McKeough et al., 2013).

Although direct improvement on cognitive tasks are often seen following direct strategy training, performance improvement rarely generalizes beyond tasks that correspond directly to taught strategies, such that so called "far transfer" is typically poor across the lifespan (Noack et al., 2009; Sala & Gobet, 2019; Simons et al., 2016; Zelinski et al., 2011). Transfer is defined as when learning in one context improves performance in a different context (Salomon & Perkins, 1989; Willis & Schaie, 2009). Transfer is a critical issue in many large-scale trials targeting cognitive constructs, as transfer may have important implications for the maintenance of independent functioning of older adults (Willis & Schaie, 2009). The core issue is that training may be seen as having little real-world value if there is not improvement beyond the target(s) of training.

Multiple review papers of cognitive intervention studies have concluded that the majority of studies have failed to detect generalized improvements in performance, with limited evidence of transfer to untrained tasks (Noack et al., 2009; Stine-Morrow & Basak, 2011). Additionally, the same interdisciplinary, international group of experts who evaluated the quality of evidence regarding cognitive training also concluded that there was no evidence of generalized cognitive improvement from specific cognitive interventions (Livingston et al., 2020). Further, a monograph examining whether intellectual, physical, and social activities improve cognitive performance across the adult life-span found that cognitive training can improve cognitive functioning, albeit with limited transfer, but with the possible exception of studies providing structured experiences in situations that demand executive coordination of skills (Hertzog et al., 2008).

Discussions regarding issues of transfer extend to the early 1900's. Woodward & Thorndike (1901) argued that the prevailing wisdom was that training generalization to untrained stimuli can only occur to the extent that the trained and untrained stimuli share features. This was coined "the law of identical elements". Others during this early period of scholarship argued that beyond identical elements, participants also needed to be consciously aware of how the trained principle might apply to the related transfer activities (Orata, 1928). Correspondingly, in the ensuing decades of training research, the overarching conclusion has been that much cognitive training shows evidence of poor transfer and generalizability of training gains.

In the 1970s, work from Paul Baltes and colleagues has found that transfer is often restricted to a narrow band surrounding the trained constructs, such that one should not expect a cognitive training program focused on improving a specific cognitive construct to generalize to other forms of cognition (Baltes & Lindenberger, 1988; Baltes & Willis, 1982). In this work, when training inductive reasoning (for example), the investigators argued that if the underlying skill of reasoning was improved, performance gains should be observed on all measures of inductive reasoning (near fluid transfer). Other measures of fluid reasoning that were related, but untrained (e.g., figural relations) might be expected to see some transfer, but less because of fewer identical elements (far fluid transfer). All other tasks would not be expected to see transfer because of the absence of identical elements (far non-fluid transfer).

Durability of cognitive training effects

Few studies have measured the durability long enough to determine the long-term effects of their cognitive training paradigm; however, encouragingly, the existing research on cognitive training interventions with longer follow-up periods demonstrate maintenance of training effects. For example, the Seattle Longitudinal Study found maintained benefits of training in inductive reasoning and spatial orientation for up to 7-years post-training (Schaie, 1993). The ACTIVE study found maintained cognitive gains from all three cognitive training groups through 5 years (Sherry L Willis et al., 2006) and inductive reasoning and speed of processing intervention groups maintained their performance advantages up to a ten-year follow-up (George W. Rebok et al., 2014).

Training effects on everyday functioning and dementia

In addition to the aforementioned long-term training effects, trained participants in ACTIVE self-reported less decline in instrumental activities of daily living (IADL’s) at ten-years (Ball et al., 2002; George W. Rebok et al., 2014; Sherry L. Willis et al., 2006). Furthermore, speed of processing and inductive reasoning trained individuals had lower rates of at-fault collision involvement, based on state-recorded motor vehicle collision data up to 6 years after study enrollment (Ball et al., 2010). Additionally, speed of processing trained participants in ACTIVE were found to be less likely to develop suspected clinical depression (Wolinsky et al., 2009) and less likely to have extensive declines in health-related quality of life (Wolinsky et al., 2006), though the possibility of confounding expectancy effects (i.e., people were aware that they were receiving training, and this may have shaped their subjective evaluations) cannot be ruled out (Simons et al., 2016).

While ACTIVE was designed to test the effectiveness of a large-scale, multi-site cognitive training intervention on cognitive health and functional independence maintenance, few studies to date have examined whether cognitive training in ACTIVE was associated with a reduced rate of incident dementia (Unverzagt et al., 2012). In one study from ACTIVE, cognitive training was not associated with a reduced rate of incident dementia after 5 years of follow-up (Unverzagt et al., 2012).

Another study from ACTIVE found that speed of processing training resulted in a reduced risk of dementia at the 10-year follow-up period, while memory and inductive reasoning training did not result in reduced dementia risk (Edwards et al., 2017). Further, each additional speed of processing training session was associated with a 10% reduced hazard for dementia (Edwards et al., 2017).

Participants were not excluded from the ACTIVE trial if they had mild cognitive impairments and had intact ADLs at baseline (Thomas et al., 2019), and one algorithmic approach using the available cognitive tests classified 30% of the baseline sample as having mild cognitive impairment (MCI) (Cook et al., 2013). Those classified as having MCI at baseline had faster self-reported functional decline and poorer performance on a performance-based measure of everyday functioning (Thomas et al., 2019; Wadley et al., 2007).

Conceptualization of cognitive training: restoration/compensation and remediation/enhancement

Rehabilitation science offers some conceptual guidance regarding how cognitive training might alter neural reorganization in the senescent brain. Rehabilitation is believed to operate through either stimulating pre-existing neural reserves, in a process known as restoration, or through compensatory-based strategies, which prompts neuroplastic reorganization to meet task demands (Sherman et al., 2017). Restoration/remediation is the process of restoring a lost function.

Rehabilitation science applies to remediating cognitive aging losses, if one conceptualization of cognitive aging is neural change and neural loss. Since it seems unlikely that behavioral cognitive training could reverse brain aging (i.e., reduced brain volume, cortical thinning), it seems reasonable to conceptualize cognitive training as compensation. Here, we define compensation as the process of learning how to develop “work arounds” for a task, function, or process. An individual is taught to change the environment, their approach, or learn to adapt their lifestyle to the difference. In cognitive training, particularly cognitive training that teaches new strategies and/or external mnemonics, such as was the case in ACTIVE (Ball et al., 2002), it is likely these strategies represent new, compensatory ways of achieving better cognitive performance.

Indeed, speculatively, it seems that inductive reasoning strategy training functions explicitly via a compensatory mechanism, such that the training may reduce the working memory demands of a task. A core component of reasoning training is to teach participants explicit strategies to group letters and recognize patterns (e.g., underlining and slashing, or adding emphasis and highlighting to increase salience of patterns). Akin to "chunking" strategies that make it easier to remember telephone numbers, these strategies reduce the need to focus on each individual letter. Thus, an important potential benefit of late-life cognitive training is that it offers compensatory strategies to help buffer against cognitive declines and preserve functioning.

Inductive Reasoning

Inductive reasoning is a specific type or category of reasoning, and is broadly defined as the ability to infer general principles from specific instances and apply these principles to novel stimuli (Saczynski et al., 2004). Inductive reasoning is focused on detecting rules or regularities and then making an inference that extends the rule or regularity (Klauer & Phye, 2008). Inductive research began nearly a hundred years ago, when Spearman determined that his *g* general intelligence factor was primarily driven by induction processes (Klauer & Phye, 2008; Spearman, 1923). Later work identified inductive reasoning as one of the more "pure" primary abilities, representing a key component of intellectual functioning or fluid intelligence (Blieszner et al., 1981; Horn, 1970, 1978; Horn & Cattell, 1967; Thurstone, 1938). Inductive reasoning has been shown to be a component of executive functioning (Lezak, 1995) and be similar to working memory ability (Salthouse, 1993), with a previous study using confirmatory factor analysis yielding consistently high correlation estimates (r = .80 to .90) between reasoning ability and working memory capacity (Kyllonen & Christal, 1990). Kyllonen and Christal suggested that the process of remembering a tried and failed rule, while searching for other rules, in an inductive reasoning task, may explain the association of inductive reasoning with working memory. Inductive reasoning has also been shown to be highly age sensitive, evincing age-related decline (Schaie, 2005).

Research has shown that inductive reasoning is very amenable to strategic instruction (Ball et al., 2002) and is associated with everyday functioning (Owsley et al., 2002; Yam et al., 2014). In studies that directly measured performance on cognitively-demanding tasks of daily living, including understanding medicine bottle labels, understanding bus schedules, and using emergency telephone information, participants with higher levels of inductive reasoning showed better everyday performance, even after adjusting for age and education (Diehl et al., 1995; Willis et al., 1992).

How is inductive reasoning trained?

As noted above, inductive reasoning training often involves strategic instruction; for example, teaching participants underlining repeated words to help focus their attention and inhibit irrelevant stimuli or drawing slashes to identify groupings. Participants are also taught dual-encoding strategies, such as speaking pattens out loud at first to also hear them while they read them, with the goal that this will later fade to an internal rehearsal strategy. Along these lines, participants are initially taught external strategies, such as marking up worksheets; however, later in the course of training, the participants are taught to try to internalize these strategies. Strategy training may also adaptively increase difficulty as progress is made; for example, exercises may become more difficult over time and more difficult strategies, such as the introduction of writing tick marks to represent skip patterns, may be introduced in later training sessions. Training also involves multiple practice exercises, with feedback, in small-group settings. In ACTIVE, inductive reasoning training was also adapted to promote real-world generalization by incorporating everyday tasks, such as finding patterns in medication schedules.

What does an inductive reasoning training study design look like?

Inductive reasoning training paradigms typically follow a pre-test post-test design with 5-10 training sessions of training in inductive reasoning between the pre-test and post-test periods, and most studies also involve a no-contact control to control for practice effects and benefit from having follow-up periods (Ball et al., 2002; Baltes et al., 1988; Baltes et al., 1989; Labouvie-Vief & Gonda, 1976; Margrett & Willis, 2006; Plemons et al., 1978; Willis & Schaie, 1986).

Inductive reasoning training effects

Training in inductive reasoning has been shown to produce durable training-related gains in inductive reasoning performance itself (Blieszner et al., 1981; Kelly et al., 2014; George W Rebok et al., 2014; George W. Rebok et al., 2014), and training in inductive reasoning has resulted in training effects of .25 to .50 standard deviation units, with the majority of participants demonstrating reliable improvement as a result of training (Schaie & Willis, 1986; Willis & Schaie, 1994).

Several earlier studies explicitly linked acquisition and use of trained strategies to improved performance. In the first published study of reasoning training with older adults, Labouvie-Vief and Gonda (1976) trained self-monitoring/metacognitive strategies to improve performance on the Inductive Reasoning task of Letter Sets (Ekstrom et al., 1976). That initial study reported that that the trained older adults improved their performance on the inductive reasoning task itself, and these training effects were maintained over a two-week period (Labouvie-Vief & Gonda, 1976).

Another study explicitly measured strategy use on inductive reasoning performance and found that participants who were trained on inductive reasoning increased their strategy use from pre- to post-test, and in turn, participants who increased their strategy use experienced larger training-related gains on inductive reasoning outcome measures. Thus, use of trained strategies was likely the important mechanism of training-related improvement (Saczynski et al., 2002), as opposed to generally improving test taking skills or merely increasing test familiarity.

Inductive reasoning training findings of transfer

In accordance with other training literature already reviewed, most of the research on reasoning training has found highly specific training effects. In one study where older adults were randomized to complete cognitive training in figural relations or a no-contact control group, while improvements were seen on near-fluid figural relations tasks, no significant differences between the training group and the control group were found on a far fluid transfer composite score of three inductive reasoning tasks (Willis et al., 1981). This was consistent with an earlier pilot study, which had included participants ranging in ages 59 to 85 years. Participants were randomized to receive figural relations training or be a part of a no-contact control group. Again, although there were significant training effects on the target of training, for the far fluid transfer measure of inductive reasoning the analyses revealed no significant training effect (Plemons et al., 1978).

Another early study explored whether a contact control (social engagement control) might show fewer differences from the treatment group than had been observed with prior no-contact control participants. Participants were randomly assigned to either inductive reasoning training or a social contact control group. Those in the inductive reasoning training group received five one-hour training sessions involving practice on new problems, trainer feedback, and group discussion of strategies (Blieszner et al., 1981). Three levels of training transfer were examined, near (inductive reasoning), far-fluid (figural relations), and far-non-fluid (other cognitive abilities). A significant training effect was found for the near fluid (i.e., direct target of training) level of transfer; however, no training effect was found for the far-fluid, nor the far-non-fluid levels (Blieszner et al., 1981).

Prior Cognitive Training Trial: ACTIVE

The largest published cognitive training intervention study, to date, has been the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) trial, which was an NIH-funded, multi-site Phase 3 clinical trial, aimed at investigating the immediate and long-term effects of cognitive training on the maintenance of cognitive functioning and independence in 2,802 community-dwelling older adults aged 65-96 years (Ball et al., 2002; Jobe et al., 2001; Sherry L. Willis et al., 2006). Participants (approximately 27% African-American and 80% female) from six field sites (University of Alabama at Birmingham, Johns Hopkins University in Baltimore MD, Wayne State University in Detroit MI, the Hebrew Rehabilitation Center for the Aged/Harvard in New England, Indiana University Purdue University Indianapolis, Pennsylvania State University in central Pennsylvania) were randomized to one of three strategy-based training arms (inductive reasoning, memory, or speed of processing), or a no-contact control. The interventions included ten 60-75 minute sessions of standardized training procedures over the course of 5-6 weeks and were conducted in small groups (Thomas et al., 2019). All three intervention arms shared common design features, including: focus on training strategies that are relevant to the ability being trained (inductive reasoning, memory, or speed of processing), modeling of strategy use by the intervention trainer, participant practice, performance feedback, the application of learned strategies to real-world contexts, and social interaction (Willis & Caskie, 2013). Across the three interventions, the first 5 sessions were primarily focused on introducing strategies and beginning to practice them, while the last 5 sessions focused on additional practice, with no new strategies introduced (Thomas et al., 2019). There was a pre-test, post-test, and follow-up (years 1, 2, 3, 5, and 10) battery consisting of measures of the cognitive training targets, everyday functioning, mood, health, health service use, and medication use.

Inductive Reasoning Training in ACTIVE: Major Findings

Inductive reasoning training in ACTIVE was focused on improving the ability to solve problems necessitating linear thinking and following a serial pattern (Jobe et al., 2001; Schaie & Willis, 1986; Willis & Caskie, 2013). As in the earlier studies, the specific implementation of inductive reasoning training in ACTIVE taught strategies for determining the pattern in a word or letter series and identifying the next item in the series (Thomas et al., 2019). Strategies included underlining repeated letters, putting slashes between series, and indicating skipped items in a series with tick marks (Willis & Caskie, 2013). These strategies were practiced in both individual and group exercises, which involved practice reasoning tasks. In a departure from earlier studies, the training also emphasized how the rules taught might apply to Instrumental Activities of Daily Living (ADLs), such as identifying a medication dosing pattern (Willis & Caskie, 2013).

As in earlier work there was near-fluid transfer. Inductive reasoning training in ACTIVE elicited improved inductive reasoning performance, a finding that was maintained through the ten-year follow-up (Ball et al., 2002; George W. Rebok et al., 2014). In a study examining who benefited from inductive reasoning training in ACTIVE, those with higher education and Mini-Mental State Exam (MMSE) scores, better health, and younger age had higher pre-test inductive reasoning performance, and higher MMSE scores were related to greater training effects and larger linear slopes over 5-years follow-up (Willis & Caskie, 2013).

ACTIVE also confirmed the strong relationship between inductive reasoning and everyday functioning. At baseline, inductive reasoning explained the most variance in three performance-based (one paper and pencil and two behavioral observation) tasks of baseline everyday functioning (Gross et al., 2011). With regard to transfer, the results were more mixed. There were no inductive reasoning training effects on performance based everyday functioning measures through ten years of follow-up (George W. Rebok et al., 2014). However, at five- and ten-years post-training, inductive reasoning-trained subjects reported significantly less functional decline, on average, in self-reported IADLs (Instrumental Activities of Daily Living; e.g. medication management, food preparation) relative to untrained controls. (George W. Rebok et al., 2014; Sherry L. Willis et al., 2006). This suggests that cognitive training may confer immediate benefit to the trained cognitive domain, but longer follow-up periods (i.e., 5-10 years) may be needed before mitigated functional decline is noticed (Thomas et al., 2019). Here, the underlying conception would be that training improved and maintained reasoning reserve capacity, but that its' everyday benefits were masked until there began to be normative functional decline (five years after training, when participants were on average 79 years of age). Then, those who had been trained were somewhat buffered against functional decline.

The caveat here is that self-reported IADLs are subjectively measured, and it is impossible to disentangle the role of unblinded participant expectancies in shaping their self ratings (Boot et al., 2013). Did participants rate themselves as more intact because this was a demand characteristic of having participated in training? Other subjective outcomes were also improved; inductive reasoning training was found to enhance internal locus of control beliefs (i.e., the sense of self-control over one's life) (Thomas et al., 2019; Wolinsky et al., 2010).

Other ACTIVE follow-up investigations have reported that (as with speed of processing trained participants), inductive reasoning-trained participants experienced lower rates of at-fault motor vehicle collision involvement following training (Ball et al., 2010). Of note, this was not a subjective report; crash rates were taken from participants' archival state motor vehicle crash records.

Collectively, these findings suggest that the cognitive training of community-dwelling older adults who are free of frank dementia may have enduring benefits. The self-reported IADL and archival motor vehicle crash findings suggest that reasoning training may have contributed to the banking of cognitive reserve, in turn contributing to the postponement of the loss of functional independence (Hartman-Stein & LaRue).

With regard to everyday transfer of inductive reasoning training, the ACTIVE study may be seen as a glass half full. Positive findings regarding self-reported IADLs and driving do suggest some delayed transfer, but relatively few transfer measures saw immediate or delayed benefit of training. Performance based measures of cognitive IADLs showed no benefit of training, despite their strong correlation with inductive reasoning performance.

One question is whether the brief dosage of training (10 sessions over five weeks) was too minimal to have a meaningful or durable impact on everyday performance. Expanding training as it was offered in ACTIVE to a higher dosage runs into feasibility concerns: Participant burden (time, travel), the need for trained tutors, and the need for spaces in which to train all make longer bouts of in-person training difficult to envision. One of the motivations for the current study, in converting inductive reasoning training to telehealth delivery, is to explore the effectiveness of training that minimizes time, travel, and space demands. While the current study per se does not propose to increase dosage over ACTIVE, if successful (i.e., users improve, and rate the training experience favorably), the telehealth platform makes it easier to imagine expanding dosages while removing some of the barriers to doing so. Also, telehealth delivery (and the digitization and careful attention to pedagogy that it requires) may represent a first step toward automating training so that, in the future, it could be adapted to be computer administered and asynchronous. The remainder of this review considers telehealth as a platform for cognitive training.

Telehealth as a Possible Means of Addressing Training Resource Limitations

Accessible Interventions are needed to support older adults (Burton & O'Connell, 2018). However, proper dementia services are less accessible to individuals in rural communities, and transportation challenges and long distances may further limit accessibility (Bédard et al., 2004). Telehealth may provide a way to make interventions to support older adults more accessible.

Telehealth defined

Telehealth, or telemedicine, is defined as the delivery of medical information or health care services from one location to another through electronic means to improve patient’s health (Organization, 2017; Tuckson et al., 2017). Early implementations of telemedicine were expensive (relying on specialized equipment and connectivity that few individuals would have in their homes). In contrast, more recent advances have allowed for inexpensive, user-friendly, real-time video interactions on consumer-grade devices, such as laptops, tablets, and smart-phones. These portable, ubiquitous devices have allowed for telemedicine to be used in a wide variety of locations, including homes (Armfield et al., 2015). One review paper summarized studies that used Skype in clinical care found that this videoconferencing tool was prevalent in the management of chronic diseases and speech and language pathology applications. In every reviewed study, aside from one, Skype was found by the authors to be feasible and beneficial (Armfield et al., 2015).

Prior telehealth work

Prior telehealth-based interventions have been found to be efficacious for cognitive behavioral therapy for insomnia and depression and anxiety (Holmqvist et al., 2014; Khatri et al., 2014). In the Holmqvist et al study, participants were given either web- (i.e., non-interactive) or telehealth-based treatment (i.e., interaction with a therapist), and participants from both interventions demonstrated significant decreases in domains of insomnia severity, fatigue, and work and social impairment (Holmqvist et al., 2014). In the Khatri et al study, online video conferencing that implemented cognitive behavioral therapy (CBT) was compared to face-to-face CBT delivery. Both groups showed reliable adherence and comparable pre-post scores of depressive symptomatology (Khatri et al., 2014). With regard to the current investigation, however, neither of these studies focused specifically on older adults. Owing to the relative recency of widespread telehealth, and due to the technological divide that may serve as a barrier to the participation of many elders in telehealth, this is an urgent area of needed research.

Not surprisingly, then, no work, to date, has been done specifically assessing the effects of telehealth cognitive-based training with older adults, which have the potential to be highly accessible, being delivered in the older adults' homes. Some initial relevant findings may be found in adjacent literature. For example, telehealth was found to be effective and feasible for rehabilitation of cognitive impairment in older adults and the development of compensatory memory skills training after stroke (Burton & O'Connell, 2018; Lawson et al., 2020). In the Burton and O'Connell study, either in-person or telehealth videoconferencing cognitive rehabilitation was delivered. Older adult participants were drawn from populations with subjective cognitive impairment, mild cognitive impairment, or dementia due to Alzheimer's disease. While the telehealth-delivered cognitive rehabilitation program appeared feasible, with high rates of participation and functional goal obtainment, the study lacked a sufficient sample size (n = 6). Moreover, two of the 6 individuals randomized to the telehealth condition withdrew prior to the start of the intervention. Telehealth was also not home-based in this study, meaning that participants in the telehealth condition completed the cognitive rehabilitation intervention in the same lab space as those in the in-person condition, just without the interventionist in the same room (videoconferencing was used) (Burton & O'Connell, 2018). In the Lawson et al study, researchers found support for their hypothesis that individual telehealth-delivered memory rehabilitation was at least comparable to face-to-face group-based methods, such that both formats were effective in improving subjective and functional memory performance in stroke patients (Lawson et al., 2020). However, this study lacked randomization, lacked a control group, and the telehealth participants were younger than those in the face-to-face condition (Lawson et al., 2020). Another study examined the feasibility and effectiveness of a videoconferencing or face-to-face cognitive intervention for community-dwelling older adults with memory problems (Poon et al., 2005). Despite a small sample size (n = 22) and lack of explanation regarding the specific content of the cognitive training intervention provided, the authors concluded that telemedicine was both feasible and acceptable for cognitive assessments and interventions for older adults with mild cognitive deficits, and found that both groups improved significantly across multiple areas of cognition (Poon et al., 2005).

Challenges of telehealth

Given generational differences in technology use and familiarity, it is easy to contemplate that there may be challenges with the home-based implementation of telehealth for older adults. In addition to technological inexperience in many elders, the perceived impersonal nature of telehealth (i.e., talking to a “machine” rather than a person) is another barrier (Demiris et al., 2009). Telehealth care may be limited by a patient’s inability to navigate the technology needed to access health services, and technological difficulties (e.g. transmission lag and network connection difficulties) may impact neuropsychological testing conducted over a telehealth medium (Hewitt et al., 2020).

Benefits of telehealth

Telehealth cognitive training interventions have the unique opportunity to potentially increase ecological validity by being done in the participants’ homes (Margrett & Willis, 2006), increase access for older adults (Demiris et al., 2009), and keep participants and trainers safe during, for example, the COVID-19 outbreak or other public health crises (Monaghesh & Hajizadeh, 2020).

Telehealth and cognitive training: Remote protocol pilot

While not strictly telehealth, one prior student conducted an at-home protocol for cognitive training in inductive reasoning. Margrett and Willis (2006) conducted a study among older adults (mean age = 71.43 years; range = 61-89 years) in which some were assigned to individual training (*n* = 30 individuals) or collaborative training (*n* = 34 individuals/17 dyads; training completed as a couple dyad), to complete ten sessions of inductive reasoning training at home or a no-treatment control (*n* = 34 individuals). Training was consistent with the ACTIVE inductive reasoning training protocol, but two modifications were made. Participants in the individual training condition also received an answer booklet, and participants in the collaborative training condition shared an answer book for each session. Additionally, the training materials for the collaborative training condition included explicit directions and pictures indicating what parts of the training material need to be completed collaboratively. Participants in both training groups were introduced to the training protocol in the first session to ensure they understood the training program. Participants undergoing individual training completed the materials individually for sessions two through ten at home, while participants undergoing collaborative training completed the training sessions with their spouse. Results found that both training groups outperformed the no-treatment control group on the Letter Series and Word Series measures at the posttest in terms of the total number of correct responses and response accuracy; additionally, the two training groups did not significantly differ from one another at the posttest on these two reasoning measures.

Overall, their findings suggested that home-based cognitive training may be viable, which bodes well for a telehealth implementation. However, a key difference is that our study utilized a telehealth delivery, while Margrett and Willis (2006) utilized a paper-and-pencil delivery. Using Zoom and Canvas came with extra challenges, including slow internet speeds, intermittent internet connectivity, and Canvas' servers going down occasionally to name a few.

Research Best Practices

In a comprehensive review of the literature on brain training programs, many of the published intervention studies were found to have major design or analytic shortcomings that precluded making definitive conclusions about training effectiveness or efficacy (Simons et al., 2016). Simons et al. proposed a set of best practices in cognitive training research help strengthen the conclusions that can be drawn (2016). Simons et al. also proposed that the ACTIVE study was a good model of best practices in most domains (except for the failure to measure participant expectancies) (2016). Indeed, the good clinical practice principles that are "baked into" large-scale clinical trials, like ACTIVE (e.g., manualization, training and certification, open archiving of protocols, measures, data and code) are principles than can assist researchers from smaller, single labs project to improve their rigor, reproducibility, and transparency (Taylor et al., 2022). The current study proposes to implement as many of these principles as possible, a practice made easier by the extensive resources available from ACTIVE.

Below, is a brief summary of the design and analytic shortcomings of many of the published cognitive training interventions that limit the evidence regarding the scientific community's ability to draw conclusions about the efficacy of cognitive training interventions. Additionally, areas where our telehealth inductive reasoning training intervention avoided these design and analytic pitfalls was provided. In instances where our intervention was not be able to avoid these potential flaws, a justification was provided.

Severe problems

Simons and colleagues (2016) classified problems with intervention studies into three distinct categories: severe, substantial, and potential. A severe problem is such that, papers with these problems should not be used to draw conclusions about the effectiveness of any intervention and preclude any conclusions about causal efficacy (Simons et al., 2016).

Severe problems include not having a pre-test baseline assessment, which means differences in post-training assessments could be due to differences at baseline, or pre-training. Our study avoided this by having a pre-training assessment to compare to post-training performance.

Additionally Simons et al. (2016) argued that not having a control group means that improvement observed from the pre- to post-training assessments could be a result of other factors than the cognitive training intervention itself. The current study utilized propensity-matching to compare our newly collected inductive reasoning trained older adult subjects to previously collected older adult participants from the in-person ACTIVE study. This propensity-matching procedure is described in more detail in the Research Design and Methods section.

The final severe problem mentioned by Simons and colleagues is that of a lack of random assignment to conditions to ensure that uncontrolled factors are equally likely to be true of the intervention and control group (Simons et al., 2016). Our study did not randomly assign participants, since we did not collect a new control group. We mitigated this concern via the aforementioned propensity-matching procedure, which served to reduce selection due to confounding variables, and thereby mimicked the benefits of an initial randomization (Dimidjian et al., 2014; Hill et al., 2021).

Substantial problems

Simons et al. (2016) also described substantial problems that would result in an intervention study only being able to provide ambiguous or inconclusive evidence for its effectiveness. Findings and conclusions drawn from studies with substantial problems, they argued, should be treated as tentative.

Key substantial problems include having a passive as opposed to an active control group, lack of pre-registration, and having a small number of participants. These potentially substantial problems were mitigated in the current study by pre-registering the study and its aims and utilizing an a priori power analysis (with a conservative estimate of effect size) to determine group sizes.

Potential problems

Simons and colleagues also described other potential problems that could limit study findings and the ability of such studies to contribute to public policy (Simons et al., 2016). Potential problems are 1) having an active but unmatched (on expectancies) control group, 2) inadequate, or vague, pre-registration, and 3) lack of blinding when using subjective outcome measures. These potential problems were mitigated by using a baseline measure of participants' expectations (the Personality in Intellectual Aging Context measure) to help match participants via propensity matching and having specific pre-registration of our study in clinicaltrials.gov (submitted after committee and

IRB approval). The current study could not blind participants and experimenters (i.e., everyone knew they received an intervention), but propensity matching participants on PIC expectancies helped mitigate this concern. The issue of blinding in cognitive training studies is challenging; the ideal placebo intervention would have identical face characteristics to the active/effective training but offer stimuli and training materials that do not produce improvement. Few cognitive training studies to date have used true placebo controls.

Telehealth Inductive Reasoning Intervention

Responding to the knowledge gap on the effectiveness of telehealth as a modality for late-life cognitive training, and responding specifically to the Simons et al. (2016) critiques, the current study was proposed. The current study sought to convert the in-person inductive reasoning training program from the ACTIVE Study to a telehealth format. Fortunately, the protocols for inductive reasoning training are highly structured and available (Taylor et al., 2022), which simplified the process of modification to a telehealth-delivered format. The study recruited 35 new community-dwelling older adults, and the study benefitted from the ability to compare telehealth-delivered training to two propensity-matched comparison groups drawn from the ACTIVE sample of 2,802, which included a sample of N = 105 no-contact control participants and N = 105 in-person inductive reasoning trained participants. The intervention had 3 aims.

Aim 1 was to determine client perceptions of and compliance with a telehealth-delivered cognitive training intervention in reasoning. It was reasoned that Aim 1 would help to determine whether an intervention of this telehealth format is feasible for older adults, which has implications for possible future, larger-scale telehealth-delivered cognitive training interventions. It was also reasoned that the study would provide helpful data regarding the participants' perceptions of the usefulness, learnability, interface quality, interaction quality, reliability, and satisfaction and likelihood of future use of this style of telehealth intervention.

Aim 2 was to determine if a telehealth cognitive training intervention in inductive reasoning is effective in improving inductive reasoning, relative to propensity-matched untrained, no-contact controls. Aim 2 would help determine whether the inductive reasoning training was effective; however, this comes with the important limitation that the in-person no-contact control group from ACTIVE received an in-person pre-training and post-training battery of cognitive measures and questionnaires, while the telehealth-delivered inductive reasoning training group was administered a battery of cognitive measures and questionnaires over telehealth, which may affect comparability.

Finally, Aim 3 was to conduct a descriptive comparison to examine whether a telehealth-delivered cognitive training intervention in inductive reasoning was comparable to propensity-matched participants from traditional face-to-face inductive reasoning training in ACTIVE. Aim 3 would help determine whether telehealth-delivered inductive reasoning training appeared to be roughly as effective as a well-validated, in-person inductive reasoning training intervention. However, this too comes with an important limitation that our sample size was underpowered to conduct an equivalency analysis. Despite that, we report the effect sizes, means, and variances needed to power a potential future, larger intervention trial.

CHAPTER 3
RESEARCH DESIGN AND METHODS

Overview

The current study was a pre-test post-test design exploring the feasibility and effectiveness of a telehealth-delivered inductive reasoning training program. The study converted the widely disseminated in-person inductive reasoning training program from a large multisite clinical trial (ACTIVE) to a telehealth-delivered format. This study also benefitted from the ability to compare telehealth-delivered training to two propensity-matched comparison groups drawn from the ACTIVE sample of 2,802 adults aged 65 and older. The current study addressed whether telehealth-delivered training could achieve inductive reasoning improvements in older adults. In addition, because the delivery of the training was novel, an important aspect of this study was to assess how telehealth-delivered cognitive training in inductive reasoning was evaluated by older adult participants in terms of usefulness, ease of use & learnability, interface quality, interaction quality, reliability, and satisfaction/anticipated future use of the telehealth intervention.

Participants

The study recruited 35 new telehealth-delivered participants to receive inductive reasoning training, and the flow of these participants from recruitment can be viewed in the consort diagram (Figure 3-1). Improvement in these participants was compared to that seen in participants from the original in-person ACTIVE trial. Therefore, there were two sets of participants in the study. The first set were the newly recruited participants for the telehealth-delivered inductive reasoning training program. The general goal of this recruitment was to mirror ACTIVE sample makeup (e.g., 76% female, 26% Black, average education of 13.5 years). The second set of participants were those drawn, via propensity-matching, from the ACTIVE trial. From the ACTIVE database, two types of participants were drawn: no-contact controls and those who received in-person inductive reasoning training. See sections titled “Propensity matched ACTIVE control sample” and “Propensity matched ACTIVE reasoning sample” for a description of participants.

Exclusion Criteria for Newly Recruited Participants

We employed the same exclusion criteria as the main ACTIVE trial to recruit participants who (a) had not yet experienced functional decline and were likely to tolerate the training protocol, and (b) were comparable to participants included in ACTIVE. The current study also introduced one new exclusion criterion beyond that used in ACTIVE, required by the mechanism of intervention delivery: participants must be shown to have telehealth readiness (discussed in detail below).

Adapting the same exclusion criteria as those in ACTIVE (Jobe et al., 2001), persons were excluded from participation if they were younger than 65 years at screening; if they had already experienced substantial cognitive decline [score of ≤22 on the Mini-Mental State Examination [MMSE] (Folstein et al., 1975)]; had a self-reported diagnosis of Alzheimer’s disease; had already experienced substantial functional decline (self-reported need for weight-bearing support or full caregiver performance of dressing, personal hygiene, or bathing 3 or more times in the previous 7 days); had medical conditions that would predispose them to imminent functional decline or death (e.g., stroke within the past 12 months, certain cancers, or current chemotherapy or radiation treatment for cancer); had recent cognitive training; were unavailable during the testing and intervention phases of the study; or had severe losses in vision (self-reported difficulty in reading newsprint), hearing (interviewer-rated, additionally, and distinct from ACTIVE, a hearing and auditory discrimination task was administered to ensure proper hearing over internet and speakers), or communicative ability (interviewer-rated) that would sufficiently impair performance to make participation impossible. As noted above, participants will also be excluded if they do not meet the criteria for telehealth readiness, see Procedures: Screening for a detailed protocol.

Recruitment Strategies

Recruitment strategies included: (a) Use of Existing Registries: the study originally recruited 35 community-dwelling, cognitively normal older adults aged 65+ from Alachua County and Jacksonville, using existing registries of older adults throughout Alachua County and Jacksonville, through the Institute on Aging at the University of Florida. The individuals in these registries had previously consented to be recruited for research studies, and Dr. Steven Anton is the Principal Investigator behind these registries and was consulted with. The Alachua County registry had approximately 3,000 community-dwelling individuals (~66% female, and ~10% member of under-represented minority groups) throughout greater Alachua County. The Jacksonville registry had approximately 2,000 community dwelling individuals (~66 female, and ~66% Black). Upon obtaining IRB approval, the Clinical Core of the Institute on Aging received the IRB approval letter and study abstract, which triggered the initiation of the process of sending out 500 IRB-approved recruitment postcards (on the study's behalf to older adults in Alachua Country. The study was briefly described on those postcards (see Appendix D) along with the minimum requirements for participation (i.e., availability and commitment). Following the postcards, 752 recruitment emails with flyers (see Appendix D) were sent to email addresses of older adults in Jacksonville. Following those 752 recruitment emails, 678 emails were re-sent to those same email addresses, removing the email addresses of individuals who had already responded to our initial email enquiry.

Other recruitment strategies employed were (b) *Participant referral*: Enrolled participants were asked if they were willing to share study flyers and postcards with individuals known to them who might be eligible and (c) Community Recruitment: Research staff contacted people 65 years of age or older for possible recruitment and shared IRB-approved recruitment flyers (see Appendix D).

Obtaining Propensity-Matched Control Groups from the ACTIVE Sample

Propensity-matching was used to attempt to reduce selection bias due to confounding variables by mimicking randomization (Dimidjian et al., 2014; Hill et al., 2021). Using propensity score matching on the dimensions of age, gender, education, race (white vs. non-white), vocabulary, word series, letter series, letter sets, Personality in Intellectual Aging Contexts factors of a) Intellectual Self-Efficacy and b) Concern about Intellectual Aging, Short-Form Health Survey factors of a) physical health and b) mental health, and baseline Mini-Mental State Exam (MMSE) scores, groups of control (available matching n = 698) and in-person inductive reasoning trained (available n = 699) comparison samples were culled from the ACTIVE database. The variables age and education were included due to known inductive reasoning baseline performance differences in the ACTIVE sample (Willis & Caskie, 2013). MMSE scores and race variables were included due to their association with larger training effects in ACTIVE (Willis & Caskie, 2013; Zahodne et al., 2015). Vocabulary, word series, letter series, and letter sets were included to further ensure that the telehealth and in-person groups are cognitively similar at baseline. The inclusion of gender and factors of intellectual self-efficacy, concern about intellectual aging, physical health, and mental as propensity-matching variables was to further ensure that the telehealth and in-person groups were similar. The MatchIt package for R (Stuart et al., 2011) was used to calculate propensity scores were calculated for each of the participants. We explored four different approaches for matching (Greifer, 2022): 1) Nearest Neighbor Matching (involves selecting the closest eligible control to be paired with the telehealth training participant; method = “nearest”, fixed matching scheme set to 3 ACTIVE participants to 1 Telehealth-trained), 2) Optimal Pair Matching (chooses matches that in totality optimize a criterion; method = “optimal”, fixed matching scheme set to 3 ACTIVE participants to 1 Telehealth-trained), 3) Optimal Full Matching (optimally assigns ACTIVE participants to Telehealth-trained participants method = “full”, variable matching scheme where we allowed the match to find between 2-5 ACTIVE participants per Telehealth-trained participant), and 4) Caliper Matching (“caliper”, variable matching scheme where matches are determined by a caliper, which is an allowed maximum difference between the Telehealth and the ACTIVE matched participants of 0.25 Standardized Mean Difference (SMD) on the aggregate composite of all the matching covariates). Each propensity matching method was compared based on Rltools (version 0.1-17) xBalance function in R, which calculates standardized mean differences along each covariate, with and without the stratification and tests for conditional independence of the treatment variable and the covariates within strata. The solution with the lowest chi-square statistic and smallest standardized mean differences was selected. This turned out to be Optimal Pair Matching for both comparison samples (Bowers et al., 2022; Hansen & Bowers, 2008).

Propensity matched ACTIVE control sample

As the SMD results reveal (see Table 3-2), all but two of the covariates (Vocabulary, Word Series) had less than a "small" effect size ($d < 0.2$) of difference in the comparison after propensity matching. Per instructions, there were 3 ACTIVE controls selected for each 1 member of the current sample, yielding $n = 105$ control participants and $n = 35$ Taylor participants at match. Participants who were not retained for the full study, $n = 4$, and their matched ACTIVE participants were dropped from further analyses, see Dropouts section for more information.

Table 3-3 shows the differences between the Taylor telehealth $n = 35$ sample and the original ACTIVE controls. Table 3-4 shows the differences between Taylor telehealth $n = 35$ sample and the propensity-matched $n = 105$ ACTIVE control subset. Table 3-5 show the improvement in balance via propensity matching. The procedure achieved exceptional covariate balance between the two groups.

Figure 3-2 shows the standardized mean differences on each variable between the total ACTIVE control sample and the Taylor Telehealth sample (pink) and for the propensity-matched ACTIVE control sample and Taylor Telehealth (blue). Of note: (1) after propensity matching all but two measures fall within 0.2 (small effect size) of a difference between the two samples, (2) values to the left of the equivalence zone were higher in the Taylor sample, while values to the right were higher in the ACTIVE sample.

Propensity matched ACTIVE reasoning sample

As the SMD results (Table 3-6) revealed, all covariates had standardized mean differences that were less than a small difference (*d* < 0.2). As with ACTIVE controls, there were 3 ACTIVE Reasoning participants selected for each 1 member of the Taylor telehealth sample, yielding *n* = 105 ACTIVE Reasoning participants and *n* = 35 Taylor telehealth participants at match. As with controls, participants who were not retained for the full study, *n* = 4, and their matched ACTIVE participants were dropped from further analyses, as described in the Dropouts section.

Table 3-7 shows the differences between the Taylor telehealth *n* = 35 sample and the original ACTIVE Reasoning sample. Table 3-8 shows the differences between the propensity matched *n* = 105 ACTIVE Reasoning subset and the Taylor telehealth *n* = 35 sample. Table 3-9 shows the improvement in balance via propensity matching. The procedure achieved exceptional (and even better than with ACTIVE controls) covariate balance between the two groups.

Figure 3-3 shows the standardized mean differences on each variable between the total ACTIVE Reasoning sample and the Taylor Telehealth sample (pink) and for the propensity-matched ACTIVE Reasoning sample and Taylor Telehealth (blue). Of note: (1) after propensity matching all measures fell within 0.2 (small effect size) of a difference between the two samples, (2) values to the left of the equivalence zone were higher in the Taylor sample, while values to the right were higher in the ACTIVE sample.

Since it appeared that the ACTIVE inductive reasoning sample appeared to match the telehealth sample better than the no-contact controls, we investigated whether there were differences between our two ACTIVE subsamples themselves. No differences were found between ACTIVE reasoning and the ACTIVE no-contact controls on each of the covariates, with the exception of Vocabulary $t(202) = -2.108$, $p = 0.036$, where ACTIVE controls outperformed in-person Reasoning participants significantly. Of note, this was a situation where ACTIVE controls were actually a better match to the telehealth trained group than ACTIVE in-person trained participants. The Vocabulary mean score at pretest for the telehealth trained group (13.52) was actually more similar to the score for the ACTIVE no-contact controls (13.56) than the ACTIVE in-person trained (12.46).

A Priori Power and Sample Size Considerations

A total effective sample size of 81 total participants, 27 of whom were newly collected telehealth participants and 54 of whom between the 2 propensity-matched control groups ($n = 27$ each), was planned. The comparison samples were propensity-matched from the existing ACTIVE no-contact controls or in-person inductive reasoning trained individuals from ACTIVE. The sample size needed was explored using G*Power 3.1.9.7 (Faul et al., 2009). We used the lower bound of the 95% CI around the Cohen's *d* value (lower bound=0.86, Cohen's *d*=.98), representing the standardized mean difference, between the in-person reasoning trained and the no-contact controls from ACTIVE (Ball et al., 2002). The dependent variable was the pre- to post-training change in a Blom-normalized reasoning composite (Blom, 1958). We computed the sample size needed for the two focal contrasts. Contrast 1 would compare telehealth-trained participants in the current study to no-contact control participants from ACTIVE. Contrast 2 would compare telehealth-trained participants in the current study to participants who received in-person inductive reasoning training in ACTIVE. Under an expected alpha of 0.025 (.05/2, because we are conducting two separate planned contrasts), minimum desired power of 0.80, and with the aforementioned lower bound of the 95% CI around the Cohen's *d* of 0.86, the design called for 27 newly recruited individuals in the telehealth-trained group. We planned to oversample by 25% (0.25 x 27 = 6.75, which rounds up to 7), so we planned to recruit 7 additional participants, to bring the recruited sample size to n = 34 (in actuality, we ended up recruiting 35 individuals). The expectation was that this would maintain the minimum necessary sample size (n = 27) at the end of the study, which was confirmed and exceeded (final compliant N with complete data was n = 31).

Dropouts

Four individuals dropped out of the study, bringing the total number of participants from 35 to 31. One individual dropped out after the pre-training assessment, but before any of the training sessions, due to personal commitments. One individual dropped out of the training after 1 session due to the commitment not aligning with their schedule. Another individual dropped out after 2 training sessions, citing difficulties with annotating on Canvas. The last individual was asked to be removed from the study because their data would not save in the Canvas platform accurately.

Design

The study design was a pre-test/post-test study, involving three-groups, a newly recruited training group that receives telehealth-delivered inductive reasoning training, and two comparator groups (a no-contact control group and an in-person inductive reasoning trained group) from the ACTIVE study that were drawn via propensity matching.

Procedures

Consenting Procedures

The sequence and flow of the study was as follows (see **Figure 3-1**). Prospective participants called our UF laboratory telephone number or emailed the study P.I. in response to recruitment materials. After taking initial contact information (name, phone number, and email address), a telehealth-administered consent, screening, and pre-training assessment was scheduled. Participants also received a HIPAA-compliant Zoom link for this initial visit via email. (Reminders of this visit were sent 24 hours and 1 hour prior to the visit, by email and text message when permitted by the participant).

At the scheduled visit, the participants received a link to the consent form (link was shared by Zoom chat, but also by email, to ensure receipt) which is hosted in REDCap (the University of Florida CTSI HIPAA-compliant cloud-hosted database portal). Following procedures outlined by the University IRB, (https://www.ctsi.ufl.edu/files/2018/01/eConsenting-using-REDCap-Instructions-updated.pdf), using screen share, the participant was shown the consent form, and each element of the consent form was discussed and summarized. The planned study, as well as its potential risks and benefits, were discussed at length, and participants had the opportunity to have any of their questions answered. At this point, participants were asked to e-sign the consent form (with finger, trackpad, or mouse, depending on their equipment) and to submit the form. Submission automatically placed the signed consent form into a file for them on REDCap. Placing the consent form as the first activity (even prior to screening) achieved several objectives: (a) screening data could be used and analyzed and (b) engaging in e-consent required the ability to turn on computer, access links, engage with a pdf, and communicate via Zoom, thus serving as an initial screening for the ability to manage technical demands of the study. Participants who could not manage the technical requirements of consent, or who were not eligible or interested based on the consent form, were not further screened.

Upon the signing of the consent form, screening questions were typically asked in the same visit. If a participant was busy, a rescheduled date was determined to finish the screening questions. A screening form (adapted from ACTIVE) was administered as an interviewer-administered questionnaire, administered over the Zoom screen share feature. Individuals who passed screening were then asked to proceed to the pre-test assessment. Individuals who met one or more exclusion criteria received the inclusion script: “This concludes your participation in the study. Thank you for answering these questions. This has been very helpful. Based on our interview today, you are not eligible to participate in the study at this time. We appreciate the time you have spent answering these questions. We will record why you screened out and provide summary data without any information that could identify you.” Additionally, per IRB guidance, ineligible participants were also asked whether we could keep the information we asked of them today, or whether they now withdrew this information. The trainers initialed their name on REDCap next to whether the participant allows us to keep their data or whether they choose to withdraw their information.

For eligible and interested participants, pre-test assessment followed immediately after screening, in the same session (an initial two-hour Zoom appointment including consent, screening, and pretesting). For participants who needed more time, or who were unwilling to continue with testing on the same day, the pre-test session was scheduled for another time. During the pre-testing, participants were asked for their availability for scheduling the training sessions with participants with similar availabilities (participants were told to expect an email in the coming weeks with their training schedule and to plan for two 1.5-hour training sessions each week, for five weeks). Participants received reminders 24-hours and 1-hour prior to each training session.

Screening and Telehealth Readiness Assessment

The first part of the screening session followed the ACTIVE screening protocol and asked participants for information about demographics and personal data (including age, education, gender, race, hearing handicap, self-rated vision, and medical comorbidities). Additionally, a hearing screener from Marra and colleagues (2020) was conducted to ensure patients were able to clearly hear and discern what the training staff were saying. Screening also included the Mini-Mental State Examination (MMSE) (Folstein et al., 1975) administered via Zoom, following telehealth administration protocols (Cullum et al., 2014; Grosch et al., 2015). All MMSE items were the same as traditional in-person paper-and-pencil MMSE. A fillable PDF was used for scoring, and participants were asked to hold up their drawings to the camera so the examiner could view them clearly for scoring. If a participant remained eligible, the participant received a telehealth readiness screening protocol. A copy of the telehealth readiness questions is included in Appendix A.

The full telehealth readiness form and technical support protocol may be found in Appendix A. If a prospective participant was unable to meet the technical specifications required (having internet of sufficient speed and having a computer with audio and visual capacity) or was unable to demonstrate a reasonable level of understanding of the programs (e.g., Zoom and Canvas), as defined by the criteria of the telehealth readiness form (Appendix A), then they were excluded from the study.

Pre-Testing Procedures

Prospective participants who were still eligible following initial screenings were asked if they wished to complete the pre-test appointment at that time or if they wished to schedule the pre-test assessment battery for another day. This battery was also conducted over Zoom. The pre-test battery consisted of measures of cognitive functioning, including inductive reasoning measures of near-transfer, language measures of far-transfer, and questionnaires assessing self-reported general health, as well as self-efficacy for intellectual aging and concern about intellectual aging (for more information on the assessment battery, see Measures section).

Intervention Procedures

Each participant was involved in the study for 7-9 weeks. The very first consenting visit was February 28th, 2022, and the very last post-testing visit occurred on May 7th, 2022. In between those two anchors, the participants were assigned to training groups on a rolling basis, as soon as a group of 5 participants with overlapping availabilities (and availabilities that worked for the trainer and at least one undergraduate research assistant) were matched to a training time. Pre-testing began on March 3rd, 2022 and concluded on March 29th, 2022. Training sessions began on a rolling basis, with the first training session 1 occurring on March 25th, 2022 and the final training session 1 occurring on April 4th, 2022. Training groups subsequently ended (completed training session 10) after 5 weeks of training, ranging from April 25th to May 4th, 2022. Post-testing began on April 26th and concluded on May 7th, 2022. A nonspecific participant visitation schedule is outlined in the timeline in Table 3-10.

All the participants completed the training program in the Spring 2022 semester. There were seven separate groups of 3-6 participants each. The groups met at regular times, over Zoom, twice a week. The training schedule for the seven groups can be seen in Table 3-11. Of note, having seven parallel groups also made it easier to reschedule participants for make-up sessions if needed (a procedure also used in ACTIVE). Make-up sessions were done on an as-needed basis, with first preference being for a participant to make-up their training session with another training group and second preference being for the participant and trainer to complete a make-up session during a time that worked well for their respective schedules. 10 standalone make-up sessions were conducted for 8 participants in total. There were 35 instances of participants moving from one session to another.

Description of inductive reasoning training

Telehealth-delivered inductive reasoning training was adapted from the in-person inductive reasoning training program used in the ACTIVE study. As in ACTIVE, training focused on improving the ability to solve problems that require linear thinking and that follow a serial pattern or sequence (Jobe et al., 2001). Participants were taught strategies to identify patterns to solve problems. These problems involved identifying the pattern in series of numbers and letters, or recognizing patterns in everyday activities, like dosing for medications. Training consisted of ten training sessions, over 5 weeks, and was conducted over Zoom. Each training session was 60-120 minutes long and typically consisted of (a) 10 minutes of introductory training exercises for basic mental abilities, such as finding patterns in schedules, (b) training exercises for everyday tasks, such as filling out medication charts, recycling charts, and understanding medicine bottle labels, and (c) a 20-question practice assessment. The first five sessions were primarily devoted to introducing strategies, such as "**scan** or **look** at every word in the schedule", "**underline** repeated days or times", "**read aloud** the schedule", and "**make slashes** between repetition of the pattern" (see Figure 3-4). The last 5 sessions were devoted to practicing their previously learned strategies. In addition, to promote generalizability to everyday functioning, the training sessions involved finding pattens in everyday activities, such as weekly schedules, medication schedules, recycling schedules, or television schedules (see Figure 3-5). For a breakdown of the content, number of practice exercises, and everyday generalizability of each session, see Table 3-12.

During and immediately following each of the ten inductive reasoning training sessions, a training summary data form was completed for each participant (see Appendix C for a copy) in REDCap. This included inductive reasoning process scores from the end-of-session, 20-question inductive reasoning assessment, the examiner's perception of the difficulty of the environment (e.g. any distractions noted, did they have difficulty hearing/seeing on occasion, or did their internet temporarily disconnect etc…), trainer subjective perceptions of whether the participant demonstrated improvement, the scheduled date of the session, the actual date of the session, the actual time the session began, and the length of the session (exposure duration).

Training program modifications for telehealth administration

Inductive reasoning training was accessed over Canvas, a learning management system (LMS) designed for course management that supports online teaching and learning. The 10-session inductive reasoning training program was converted from its paper-pencil format to an online, Canvas course format, with all participants registered for the course. The UF Quick Registration System, aka "QuickReg", is an online application system that allowed the participants to enroll into the Canvas-delivered, non-college credit training program, and participants enrolled prior to the first training session, during the pretest. The first session was approximately an extra 30 minutes long to work through Canvas and Zoom technical difficulties that arose. Participants were provided instruction on Canvas explaining the Canvas and Zoom features necessary for training during the pretest, but many required frequent reminders during the first training session. Email reminders were sent to the participants 24 hours and 1 hour before each of the training sessions, and these emails included a Zoom link, a Canvas link, and instructions for navigating to the “Reasoning Project Site” (the Canvas “course” with the training program. Each of the 10 inductive reasoning training sessions from the original, in-person ACTIVE study was converted to a separate module in Canvas.

Within each of the 10 modules, there were sequentially numbered links to portions of training that must be completed in succession before the next link was to be accessed. Canvas allows the didactic content (“pages”) and practice exercises (“annotated assignments”) to be ordered in a fixed order for all participants, so participants must achieve milestones before they can progress from one activity to the next. However, this feature was not utilized in case participants needed to skip a training session or in case an activity would not load, preventing participants from completing the trainings. The trainer slightly modified the ACTIVE inductive reasoning trainer manual’s script to fit the telehealth format of this intervention. The trainer translated the script into an easier to process Word document and kept that document on a separate monitor to ensure that the participants were on the right activity and that they were proceeding within the time frame allotted for each activity. The scripts (see Appendix E) included the name of each activity in Canvas, the time – in minutes – for each activity, and a script to guide the trainer. This allowed for standardized administration across the training groups and consistent administration to the ACTIVE inductive reasoning training program, with only slightly modifications appropriate for the Zoom and Canvas mediums. Timing guides helped ensure that all the participants progressed through the training sessions at a similar pace, and that the sessions did not run too long.

For tasks that required the participants and instructors to underline, draw slashes, or highlight text etc… the “annotated assignments” function in Canvas was used. An example of the **underline** and **make slashes** strategies applied to an example Letter Series problem (Thurstone & Thurstone, 1947) can be seen in Figure 3-6. This demonstrates how the participant used the paintbrush feature in Canvas to underline repeated letters or words, make slashes between groups of letters or words, make tick marks indicating skipped letters, and also circle answer choices. Each session was concluded with a 20-question inductive reasoning quiz, also on Canvas, to measure training process data that informs on participants’ progress and learning trajectories from session to session. The final test is described in greater detail in the **Training Summary Data Form & Within Training Session Measures of Inductive Reasoning Performance** section of this document.

To enhance the rapport between the participants and the instructors, casual conversations were allowed to occur while waiting for everyone to log in to the Zoom room and Canvas.

Intervention team roles

Primary interventionist: Telehealth inductive reasoning training sessions were primarily led by the principal investigator (BT), a doctoral trainee in clinical neuropsychology. BT had experiencing completing neuropsychological evaluations in-person and over telehealth. BT also had experiencing administering therapy services via telehealth and using the Zoom interface.

Intervention Certifier: Dr. Michael Marsiske has extensive background in cognitive training interventions, including his role as one of the Principal Investigators for ACTIVE. He helped train Brad. Dr. Marsiske also has a background of using Canvas for his courses and was able to assist Brad in the creation of a Canvas course for this intervention.

Backup interventionists: Mr. Joshua Owens played a key support role for Brad by serving as a backup intervention leader when Brad was unavailable for one of the sessions and helping assist older adults in breakout rooms for two other training sessions as part of his practice for being a backup intervention leader.

Training assistants: There were 9 total Research Assistants assisting with this study. One to two research assistants were assigned to each group of participants. These individuals were occasionally sent to a breakout room with a participant, periodically throughout each training session, if additional one-on-one technical support was needed or to “catch up” to the rest of the training group. The small-group leaders/research assistants were there to help with group exercises in breakout rooms. Additionally, Ryan Faulkner led one training session with the assistance of Madison Verdone. Both research assistants received scaffolded training in leading training sessions under the live supervision of Brad before this training session.

Training and certification procedures for intervention team members

As was the case in ACTIVE, the trainer and data collectors underwent certification prior to administering training sessions. However, for being certified as a trainer, the same protocol was not followed as was used in ACTIVE. Brad Taylor first reviewed standard study procedures and. practiced with a mock participant, followed by one older non-participant volunteer. Two of the training sessions were video-recorded, and a certified trainer (MM) reviewed the sessions and provided session-by-session feedback. The certification procedures from ACTIVE’s manual of operations were not strictly followed (https://osf.io/nyad9/).

Additionally, the Research Assistants participated in meetings prior to the training protocol commencing. The team also helped design the Canvas page for the intervention. Further, the team was shown the Zoom features that were used during training and practiced working through inductive reasoning problems together.

Ongoing quality monitoring and fidelity checks

The supervisor, Dr. Michael Marsiske (MM), was responsible for monitoring training activities completed over Zoom. After certification, two sessions were reviewed by MM, asynchronously using recordings. Reviews were followed by feedback sessions.

The small-group leaders were monitored by Brad. Brad met with the research assistants after each training session, and as a part of that meeting, provided them feedback regarding their performance.

The research assistants had weekly Zoom meetings with Brad Taylor for general feedback and preparation for the next sessions. The research assistants were encouraged to share their experiences so other research assistants could learn from them. Brad praised the small-group leaders for the helpful behaviors he noted and provided constructive feedback for ways they may improve their behaviors for future sessions. Additionally, time was spent going over the content of the next sessions and what specific techniques they needed to practice to be ready for their upcoming sessions with the participants.

Post-Testing Procedures

After the telehealth-delivered inductive reasoning training period, the participants completed the post-test assessment, which included the same measures of cognitive functioning, including inductive reasoning measures of near-transfer, a language measure of far-transfer, and a questionnaire assessing their subjective evaluation of their self-efficacy for cognitive aging and their concern about aging. Additionally, another questionnaire, a modified version of the Telehealth-Usability Questionnaire (TUQ) (Parmanto et al., 2016) (instrument to assess patient perceptions of medical telehealth), assessing their perceptions of the usefulness, ease of use & learnability, interface quality, interaction quality, reliability, and satisfaction & future use of the telehealth intervention will be administered. See Appendix B for the telehealth assessment questionnaire. The full procedural sequence flow diagram can be found in Figure 3-7. After participating in the study, participants were paid $50.

Participation Retention

Prior communication to reduce dropout

Before participants committed to participating in the study, the time commitment required to take part in the study, including the pre- and post-training assessments and the inductive reasoning telehealth training program, was clearly articulated to them during the screening and informed consent processes. Additionally, the study’s training sessions dates were provided to each participant prior to the first training session.

Retention strategies after enrollment

Throughout the intervention, several retention strategies were employed to reduce sample loss, including offering make-up sessions to participants who were unavailable for their regularly scheduled training session. Additionally, sessions were scheduled for times when participants and study staff were available. This was accomplished by using a deidentified When2Meet, to find overlapping availabilities of research assistants and participants. Adherence to inductive reasoning training sessions was systematically monitored and recorded with a REDCap session form that was updated on a session-by-session basis and reviewed after each session with study staff. The participants were emailed a recurrent Zoom link and Canvas link before the first training session and were sent the Zoom link again 24 hours and 1 hour before each training session. Occasionally, in rare instances of network connection problems, participants were called so they may still hear the trainer over the phone. If a participant Indicated their inability to attend their scheduled upcoming training session, then we attempted to schedule for a make-up session with another training group or for another time for a one-on-one make-up session. Attendance, including the time range (on time, early, or late) participants joined the session and left, was monitored, and recorded in a spreadsheet for each training session. Participants who did not arrive within 5 minutes were called and reminded of the session. “No-shows” were called after their scheduled training session concluded and asked if they would like to attend another group’s training session as a make-up session.

Data Entry and Data Management Procedures

Self-administered forms: maintaining data quality

In addition to the e-consent form, which uses REDCap, many of the pre-test and post-test cognitive tests and questionnaires (e.g., Short-Form Health Survey, Personality in Intellectual-Aging Contexts, Word Series, Letter Series, Letter Sets, Vocabulary, and Modified Telehealth Usability Questionnaire) were programmed to be administered directly in REDCap, such that when a participant was finished, it was automatically saved to their file in REDCap. Data entry errors were avoided by having entry fields codes, such that only valid values could be entered (e.g., a Likert scale that could only accept values 1-7, or a phone number needing to be 7 digits). Additionally, forced response (i.e., a form could not be submitted if an item was skipped) was used to reduce missing data.

Electronic data entry for all forms

The measures (e.g., Demographics and Personal Data; Telehealth Readiness Form; MMSE; Vocabulary; Inductive Reasoning scores for Word Series, Letter Series, and Letter Sets; the Training Summary Data Form; the Short-Form Health Survey; Personality in Intellectual-Aging Contexts questionnaire; and Modified Telehealth Usability Questionnaire) were made into electronic REDCap forms that were filled out in real time while completing the interview, thus removing the need for any paper forms in this study. Some questionnaires (the Short-Form Health Survey, Personality in Intellectual-Aging Contexts questionnaire, and Modified Telehealth Usability Questionnaire) were sent to participants to complete on their own time soon after the pretest or posttest visits.

Data Security

The data was stored in password-protected, encrypted databases (REDCap & Canvas), which could only be accessed by the study staff. These existed on a secured, password-protected, and HIPAA-compliant server maintained by the College of Public Health and Health Professions division of information technology (IT).

Merging and storing final de-identified data set

The College of Health and Health Professions maintains secure cloud-based storage folders for all credentialed college personnel. To gain and maintain access, personnel must show evidence of valid annually renewed HIPAA and local IRB training. Access requires authentication through the University of Florida's Shibboleth-based authentication server and uses dual factor authentication (password plus a push message to a secondary device) to enable access. All study materials were stored in a project folder (p:\ drive, in the College drive labeling system). Access to the folder was restricted to authenticated users on the project. Access will be removed upon personnel termination from the project. It is on this drive that downloading data from REDCap occurred. Missing data forms were identified in real time by having REDCap flag missing values with a color-coding scheme (e.g., REDCap turns the form red on its dashboard. When a form was missing, this triggered project staff to query participants to obtain the form, reducing missing data.

Measures

See Table 3-13 for all the measures.

Demographics and Personal Data

ACTIVE's telephone screener was used to assess demographics and other personal data, including age, education, gender, race, hearing handicap, self-rated vision, and medical comorbidities. The rationale for including these items was to use them for excluding ineligible participants and for propensity-matching participants to be similar in age, education, gender, and race. They also enabled us to characterize the sample descriptively.

Telehealth Readiness

As noted above, the rationale for including this measure was to screen out individuals for whom the telehealth challenges of participating in assessment and intervention might be so burdensome, that they could not be expected to have satisfying participation in training. A telehealth readiness and technical support form was used to assess whether the prospective participants were ready for telehealth-delivered cognitive training and for providing technical support for accessing Zoom and Canvas; see Appendix A. A hearing screener from Marra and colleagues (2020) was conducted to ensure patients were able to clearly hear and discern what the training staff were saying. Six word-pairs (one red and one blue) that were phonetically difficult to distinguish were simultaneously presented on the screen, via a PowerPoint, and one of the words was read to the participant. The participant was then be asked to identify the color of the word that was spoken by the examiner. If 2 or more pairs of words were answered incorrectly by the participant, we asked them to adjust their volume and try the incorrect word pairs again.

Mental Status/Cognitive Screening

The MMSE was developed as a brief screening measure of impaired cognition (Cockrell & Folstein, 2002). The total score ranges from 0-30, representing the summed score of individual items, and indicating the current severity of cognitive impairment (Cockrell & Folstein, 2002). The rationale for including this measure was principally exclusion of individuals below threshold; the measure was also used to describe the sample and in propensity matching. The Mini-Mental State Examination (Folstein et al., 1975) was conducted over Zoom as a measure of general cognition and for exclusionary purposes if participants scored below 23/30 (none of the screened participants scored below 23). This cutoff was chosen to match that used in the ACTIVE Study. In ACTIVE, screeners had to administer, score, and make an exclusion decision in real time, and based on MMSE performance, decide whether to administer the next measure or not. Thus, rather than age- and education-referenced normative cutoffs, ACTIVE picked a global cut score which is set low enough to exclude very few people on the basis of advanced age and low education. To help ensure a comparable sample, we followed the same procedure; one rationale from the ACTIVE study for this cutoff was not to exclude individuals with fewer years of education or members of underrepresented minority groups. Test-retest reliability estimates from relatively short test-retest intervals (< 6 months) usually range from 0.80 to 0.95 (Tombaugh & McIntyre, 1992).

Subjective Health

The Short-Form Health Survey (SF36) (Ware Jr & Sherbourne, 1992) was included for (a) use in propensity matching, and (b) use in characterizing the sample. Participants were asked to rate the domains of general health, vitality, role limitations due to emotional problems, general mental health, bodily pain, social functioning, role limitations due to physical problems, and physical functioning across 36 total questions. This measure was constructed to survey health status, and scores range from 0-100 for each of the domains, with higher scores representing a more desirable health state (https://orthotoolkit.com/sf-36/). To create two aggregated component scores (physical & mental component scores), the 8 individual scales were Z-scored, according to published norms, and then weighted sums based on published factor scoring coefficients were calculated (Ware et al., 1994). The Physical Component Scale (PCS) and Mental Component Scale (MCS) were constructed to reduce the eight-scale profile to two summary measures, without substantial loss of information. The PCS and MCS have a high reliability (0.93 PCS and 0.88 MCS). The SF36 as a whole has been found to have internal consistency reliability ranging from 0.83 to 0.93 for the eight scales (Gandek et al., 2004) and test-retest reliability for all eight SF36 scales ranged from 0.648 to 0.868 after 4 weeks (Andresen et al., 1996). The SF36 has been found to be sensitive to change (Busija et al., 2008; Jenkinson et al., 1995).

Inductive Reasoning

As inductive reasoning is the target of training, we measured it with multiple instruments to assess near transfer. Three measures of inductive reasoning were employed at both the pre-training and post-training visits [Word Series (Gonda & Abilities, 1985); Letter Series (Thurstone & Thurstone, 1947); and Letter Sets (Ekstrom et al., 1976)]. These three measures are standardized, timed assessments (Jobe et al., 2001), and the same items were presented at both pretest and posttest. Measures were converted from their paper-pencil format to telehealth-delivered Canvas formats. Across previous intervention trials, including ACTIVE and the Seattle Longitudinal Study, these measures have been found to be sensitive to change (Ball et al., 2002; Saczynski et al., 2002; Schaie, 1993).

The Word Series task consists of a series of words (days of the week, months of the year) in a pattern (Gonda & Abilities, 1985). Participants were asked to select the subsequent day or month that should come next in the series, from a choice of five options. The participant was given 30 sequences to complete in 6 minutes, with scores ranging from 0-30 (1 point per each correct item), and higher scores represented better performance. Word Series has a test-retest reliability of 0.84 (Ball et al., 2002).

The Letter Series tasks consists of a sequence of letters with an identifiable pattern, and the participants were asked to select the letter that should come next in the series from a choice of five letters (Thurstone & Thurstone, 1947). The participant was given 30 sequences to complete in 6 minutes, with scores ranging from 0-30 (1 point per each correct item), and higher scores represented better performance. Letter Series has a test-retest reliability of 0.86 (Ball et al., 2002).

The Letter Sets task consists of five sets of four letters each (Ekstrom et al., 1976). Participants were asked to determine which set of letters within the five is different from the others. Participants were given 15 items to complete in 7 minutes. Scores range from 0-15, with one point given for each correct response, and higher scores represented better performance. Letter Sets has a test-retest reliability of 0.69 (Ball et al., 2002).

The remaining cognitive measure in this study was administered as a measure of far non-fluid transfer. This is a measure for which we did not expect a training effect; its inclusion thus was meant to confirm that training effects were specific to the training target and did not reflect generalized improvement in test taking skills or other non-specific performance factors.

Language

Vocabulary was assessed as a non-fluid transfer measure. The Vocabulary (Ekstrom et al., 1976) test assesses recognition vocabulary. The measure is sensitive to change (Lachman & Leff, 1989), and has a split-half reliability of 0.88 and 5-year test-retest reliability of 0.78 (Lachman & Leff, 1989). Participants were asked to select a word with a meaning similar to an identified word. Participants were shown a list of 18 target words, each of which has five response choices. Participants were tasked with picking the word that is closest in meaning (i.e., synonym). This task had a 4-minute time limit for participants to complete the test. Scores reflected the total number correct and range from 0-18.

It was not possible to directly compare ACTIVE and Taylor Telehealth participants in the same timeframe, because ACTIVE did not re-administer Vocabulary until the 1-year posttest (after about 50% of Inductive Reasoning-training participants had also received four booster sessions). We chose to compare our sample the ACTIVE groups despite this, since otherwise there was no way for us to determine if any change that would be observed in the Taylor Telehealth sample was only retest-related change, or if it was substantially different from what was observed in ACTIVE. We discuss this limitation of unequal time frames in the Discussion section.

Self-Efficacy and Concern About Intellectual Aging

The second transfer measure is Self-Efficacy and Concern About Intellectual Aging, as assessed by the Personality in Intellectual-Aging Contexts (PIC) measure (Lachman et al., 1982). The rationale for inclusion of this measure was both for propensity matching and as a measure of pre-post change in self-rated beliefs about cognitive aging. Participants were given the PIC measure (Lachman et al., 1982); the 36-item short form consisted of six scales, all representing domain-specific evaluations about one’s cognitive aging. The six scales (Internal control, Powerful Others control, Chance control, Morale, Achievement Motivation, Anxiety) assessed subjective beliefs about one’s cognition, and can be summarized as two higher order factors: (a) Intellectual Self-Efficacy; (b) Concern about Intellectual Aging. The items were presented in a 6-point Likert response format with six anchors: strongly agree, agree, slightly agree, slightly disagree, disagree, and strongly disagree. The test-retest reliability of the PIC measure is fairly high, ranging from 0.74 to 0.88 (Lachman et al., 1982). Scores range from 12 to 72 for Intellectual Self-Efficacy and 24-144 for Concern about Intellectual Aging. As with Vocabulary, ACTIVE administered the follow-up PIC scales at a different time point (1-year follow-up), and thus the comparison is inexact, but is used to provide a preliminary indication of whether any changes observed in the Taylor Telehealth sample were unique or different from what was seen in ACTIVE.

Training Summary Data Form & Within Training Session Measures of Inductive Reasoning Performance

On a session-to-session basis, to measure participants' inductive reasoning gains throughout training, a brief end-of-session inductive reasoning assessment was conducted at the conclusion of each training session. Just like the rest of the training protocol, this was completed on Canvas. At the end of each respective training session on Canvas, the participants clicked on a link within their module (10 modules, 1 per training session) that directed them to a Canvas-delivered inductive reasoning test. This 20-question end-of-session inductive reasoning task was designed to mimic the format of the in-person, paper-pencil inductive reasoning task from ACTIVE, albeit with increased spacing to allow for easier annotating and circling of answers for an online format. The answer choices were multiple choice. See Table 3-13 for the tests and their targeted cognitive domains.

During and immediately following each of the ten inductive reasoning training sessions, a training summary data form (see Appendix C for the ACTIVE version and the Telehealth version modeled after the ACTIVE version) was completed for each participant in REDCap. This included inductive reasoning process scores from the end-of-session 20-question inductive reasoning assessment. Additional items on the form measured exposure, adherence, interruptions, perceived effort, but none of these variables were used in the current study.

Post-hoc coding of strategy use at posttest

After completion of the planned analyses, upon noting some differences in improvement patterns in the current study compared to the ACTIVE Inductive Reasoning group, we examined all participants' posttest strategy use. This enabled us to answer the question of whether those who used strategies (e.g., underlines, slashes, etc.) differed in their pre-post change from those who did not. We operationally defined participants as strategy users if there was evidence of strategy use on one or more items of each measure. Thus, three variables were derived: Word Series Strategy Use, Letter Series Strategy Use, and Letter Sets Strategy Use. For each dichotomous variable, 0 = No strategy use, 1 = one or more items showed strategy use.

Telehealth Usability

Telehealth usability was assessed at post-test, using a modified version of the Telehealth-Usability Questionnaire (TUQ) (Parmanto et al., 2016) (instrument to assess patient perceptions of medical telehealth), designed to measure dimensions of usefulness, ease of use & learnability, interface quality, interaction quality, reliability, and satisfaction & future use. The original TUQ demonstrated good to excellent internal consistency on the dimensions of usefulness, ease of use, effectiveness, reliability, and satisfaction, with raw Cronbach Coefficient Alpha values ranging from 0.79 to 0.92 (Parmanto et al., 2016). The content validity of the TUQ has also been demonstrated in other studies (Daihua et al., 2015; Yu et al., 2017). This helped to determine whether a cognitive training intervention in inductive reasoning of this telehealth format is feasible for older adults, which has implications for possible future, larger-scale telehealth-delivered cognitive training interventions. It provided helpful data regarding the participants' perceptions of the usefulness, learnability, interface quality, interaction quality, reliability, and satisfaction and likelihood of future use of this style of telehealth intervention. See Appendix B for the questionnaire.

Statistical Analysis Plan

Analyses

Data inspection preparatory to analyses inspected the distributions of all variables for outliers and any incorrect values. MM and BT conducted the statistical analyses, which were performed using statistical software, R and SPSS. Analyses that BT conducted were under the close supervision from the mentor (MM) who has taught graduate-level statistics courses.

Aim 1

To determine client perceptions of and compliance with a telehealth-delivered cognitive training intervention in inductive reasoning.

Hypothesis 1.1

It was hypothesized that subjects would provide "agree" or better ratings on questions asking for their perceptions of the usefulness, ease of use & learnability, interface quality, interaction quality, reliability, and satisfaction & future use of the telehealth inductive reasoning training program.

Analytical Approach

Descriptive statistics of participant evaluation and compliance metrics, which we examined using horizontal segmented bar charts that illustrated the degree of endorsement of each level of the item.

Hypothesis 1.2

It was hypothesized that participants would attend at least 80% (8/10) of training sessions, on average, a rate comparable to that observed in in-person trials.

Analytical Approach

A descriptive report of the percentage of participants who achieved 80% or better attendance at training sessions was obtained.

Aim 2

To determine if a telehealth cognitive training intervention in inductive reasoning was effective in improving reasoning.

Hypothesis 2.1

It was hypothesized that those who received inductive reasoning telehealth training would perform significantly better at an immediate posttest that occurred about one week after the conclusion of a 5-week, 10-session training program on measures of inductive reasoning (Word Series, Letter Series, Letter Sets). Transfer measures (Vocabulary, Self-Efficacy, Concern about Aging) were not expected to show pre-post change.

Hypothesis 2.2

Participants' training gains would exceed simple practice effects on Word Series, Letter Series and Letter Sets (but not Vocabulary, Self-Efficacy, or Concern about

Aging), as evidenced by telehealth-trained individuals showing greater gain than a matched control group culled from the Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE) clinical trial.

Analytical Approach

A two-way mixed analysis of variance was conducted for each of the four cognitive measures (Word Series, Letter Series, Letter Sets, Vocabulary) and for each of the two PIC self-evaluation measures (Self Efficacy, Concern about Aging) as outcomes with Occasion (pre-training, post-training/1-year follow-up [in the case of Vocabulary and the two PIC measures in the ACTIVE samples]) and Group (Telehealth, ACTIVE In-Person Inductive Reasoning Training, and ACTIVE Control) as the independent variables. Planned contrasts assessed whether the magnitude of pre-post change differs between (1) telehealth and ACTIVE in-person, and (2) telehealth and ACTIVE control.

Aim 3

To conduct a descriptive comparison to examine whether a telehealth cognitive training intervention in inductive reasoning was comparable to traditional face-to-face inductive reasoning training.

Hypothesis 3.1

Participants receiving telehealth training would experience training gains similar to those observed in participants who received traditional face-to-face inductive reasoning training. These results would provide us with effect size estimates to power a future equivalency trial. No differences between ACTIVE and Taylor Telehealth were expected for any of the measures (Word Series, Letter Series, Letter Sets, Vocabulary, Self-Efficacy, Concern About Aging). Thus, the null hypothesis was expected to be true.

Analytical Approach

A two-way mixed analysis of variance was conducted for each of the four cognitive measures (Word Series, Letter Series, Letter Sets, Vocabulary) and two PIC self-evaluation measures (Self Efficacy, Concern about Aging) as outcomes with Occasion (pre-training, post-training/1-year follow-up [in the case of Vocabulary and the two PIC measures in the ACTIVE samples]) and Group (Telehealth, ACTIVE In-Person Inductive Reasoning Training, and ACTIVE Control) as the independent variables. Planned contrasts assessed whether the magnitude of pre-post/1-year follow-up change differs between (1) telehealth and ACTIVE in-person, and (2) telehealth and ACTIVE control.

Because the expectation for Aim 3 was of no difference between Taylor Telehealth and ACTIVE Reasoning participants, we recognized that absence of evidence (support for the null hypothesis) did not constitute evidence of absence (i.e., no difference between groups). To make more definitive statements about the equivalence of improvement in both groups, a formal statistical test of equivalence was required. However, the current study was underpowered. Using PASS (Hintze, 2008) and data from the ACTIVE study, we expected that inductive reasoning trained participants would improve on the inductive reasoning composite (a standardized average of Word Series, Letter Series and Letter Sets) by about 2 points. We defined equivalence as the two groups being within 10% of each other (0.2) in the amount of improvement. At alpha = 0.05 and power = 0.80, the program determined that we would need 1296 participants (648 in each group), which far exceeded available resources for this study. Nonetheless, purely to investigate whether there was even trend-level evidence in support of our equivalence hypothesis, we did conduct a (grossly underpowered) equivalency analysis. The analysis used the R package TOSTER (Lakens, 2017), and defined equivalency as the two groups pre-post change being within 10% of each other and conducted two one-sided tests to test the significance of the difference between the upper and lower ends of the equivalency interval with the range defined by the 90% confidence interval around the mean difference.

Table 3-1. Chi-square statistics for the 4 propensity-matching approaches for the Telehealth vs. ACTIVE controls comparison and Telehealth vs. ACTIVE reasoning comparison

	Telehealth vs. ACTIVE controls	Telehealth vs. ACTIVE reasoning
Nearest neighbor	χ^2 (13) = 5.21, p = 0.721	χ^2 (13) = 2.28, p = 1.000
Optimal pair matching	χ^2(13) = 3.53, p = 0.995	χ^2 (13) =1.58, p = 1.000
Optimal full matching	χ^2 (13) = 9.15, p = 0.762	χ^2 (13) = 3.91, p = 0.992
Caliper matching	χ^2 (13) = 9.19, p = 0.758	χ^2 (13) = 3.36, p = 0.996

Table 3-2. Test of covariate balance under optimal pair matching, ACTIVE Controls and Taylor telehealth

Group:	ACTIVE controls		Taylor telehealth		
N:	105		35		
	Mean	*SD*	*Mean*	*SD*	*SMD*
Sex (1=male, 2 = female)	1.67	0.47	1.63	0.49	0.079
Age (years)	71	4.66	71.21	4.44	0.045
White (1=white, 0=not white)	0.86	0.35	0.83	0.38	0.078
Years of Education	15.53	2.7	15.49	1.77	0.021
MMSE	28.74	1.19	28.6	1.29	0.115
Vocabulary, # correct	13.56	3.37	12.74	3.37	0.243
Word series, # correct	11.45	5.1	10.49	3.54	0.219
Letter series, # correct	14.32	5.76	13.89	6.06	0.074
Letter sets, # correct	8.07	2.76	7.8	2.61	0.099
PIC Self Efficacy	32.55	2.84	32.11	3.22	0.144
PIC Concern About Aging	2.6	4.79	3.34	4.57	0.16
SF36 Physical Health	-0.2	0.91	-0.31	1.02	0.113
SF36 Mental Health	0.64	0.67	0.59	0.56	0.081

Notes: MMSE = Mini-mental status examination, PIC = Personality in Intellectual Aging Contexts Inventory, SF36 = 36-Item Short Form RAND Health Survey, *n* = sample size, *SD* = standard deviation, *SMD* = standardized mean difference (Cohen's d).

Table 3-3. Difference between the Taylor telehealth and original ACTIVE controls.

Summary of Balance for All Data:	Means Taylor Telehealth	Means ACTIVE Control	Std. Mean Diff.	Var. Ratio	eCDF Mean	eCDF Max
distance	0.1822	0.0427	0.9117	4.1962	0.3733	0.6336
Sex (1=male, 2 = female)	1.6286	1.7392	-0.2257	1.2448	0.0553	0.1106
Age (years)	71.2061	73.8167	-0.5875	0.5718	0.1164	0.2533
White (1=white, 0=not white)	0.8286	0.7079	0.3202	.	0.1207	0.1207
Years of Education	15.4857	13.3696	1.1943	0.4312	0.1585	0.532
MMSE	28.6	27.31	1.0016	0.4259	0.1613	0.2962
Vocabulary, # correct	12.7429	12.2265	0.1531	0.7278	0.0605	0.1415
Word series, # correct	10.4857	9.3159	0.3302	0.5366	0.0591	0.2431
Letter series, # correct	13.8857	9.6453	0.6995	1.1967	0.1533	0.3524
Letter sets, # correct	7.8	5.6617	0.8193	0.8147	0.1361	0.3521
PIC Self Efficacy	32.1143	30.4456	0.5178	0.632	0.081	0.2913
PIC Concern About Aging	3.3429	7.6954	-0.9515	0.4852	0.1606	0.3268
SF36 Physical Health	-0.3082	-0.6731	0.358	0.9652	0.1212	0.2545
SF36 Mental Health	0.5862	0.3753	0.3734	0.4348	0.0662	0.1704

Notes: distance = Mahalonobi's squared distance, a measure of multivariate distance of each case from overall centroid using all covariates, MMSE = Mini-mental status examination, PIC=Personality in Intellectual Aging Contexts Inventory, SF36 = 36-Item Short Form RAND Health Survey, Std. Mean Diff.= Standardized Mean Difference, Var. Ratio = Variance Ratio, eCDF Mean = empirical cumulative density function mean, eCDF Max = maximum empirical cumulative density function difference, Std. Pair Dist. = standardized pair distances.

Table 3-4. Difference between the Taylor telehealth and propensity-matched ACTIVE controls.

Summary of Balance for Matched Data:	Means Taylor Telehealth	Means ACTIVE Control	Std. Mean Diff.	Var. Ratio	eCDF Mean	eCDF Max	Std. Pair Dist.
distance	0.1822	0.1558	0.1722	1.3645	0.011	0.1714	0.1834
Sex (1=male, 2 = female)	1.6286	1.6667	-0.0777	1.0712	0.019	0.0381	1.1268
Age (years)	71.2061	71	0.0464	0.9079	0.0296	0.0952	1.1452
White (1=white, 0=not white)	0.8286	0.8571	-0.0758	.	0.0286	0.0286	0.7328
Years of Education	15.4857	15.5333	-0.0269	0.4318	0.0592	0.181	1.4029
MMSE	28.6	28.7429	-0.1109	1.1652	0.0226	0.0857	0.9095
Vocabulary, # correct	12.7429	13.5619	-0.2429	0.9991	0.0561	0.2667	1.0731
Word series, # correct	10.4857	11.4476	-0.2715	0.4824	0.0652	0.2286	1.449
Letter series, # correct	13.8857	14.3238	-0.0723	1.1075	0.0279	0.0762	1.0337
Letter sets, # correct	7.8	8.0667	-0.1022	0.8939	0.0214	0.0857	1.1604
PIC Self Efficacy	32.1143	32.5524	-0.136	1.2834	0.026	0.1333	1.0256
PIC Concern About Aging	3.3429	2.5952	0.1634	0.9101	0.0348	0.1524	1.1191
SF36 Physical Health	-0.3082	-0.1984	-0.1077	1.2448	0.0321	0.0952	1.0516
SF36 Mental Health	0.5862	0.6364	-0.0889	0.7056	0.0453	0.1429	1.1745

Notes: distance = Mahalonobi's squared distance, a measure of multivariate distance of each case from overall centroid using all covariates, MMSE=Mini-mental status examination, PIC=Personality In Intellectual Aging Contexts Inventory, SF36 = 36-Item Short Form RAND Health Survey, Std. Mean Diff.= Standardized Mean Difference, Var. Ratio = Variance Ratio, eCDF Mean = empirical cumulative density function mean, eCDF Max = maximum empirical cumulative density function difference, Std. Pair Dist. = standardized pair distances.

Table 3-5. Improvement in balance due to propensity matching.

Percent Balance Improvement:	Std. Mean Diff.	Var. Ratio	eCDF Mean	eCDF Max
distance	81.1	78.3	97	72.9
Sex (1=male, 2 = female)	65.6	68.6	65.6	65.6
Age (years)	92.1	82.7	74.5	62.4
White (1=white, 0=not white)	76.3	.	76.3	76.3
Years of Education	97.7	0.2	62.7	66
MMSE	88.9	82.1	86	71.1
Vocabulary, # correct	-58.6	99.7	7.3	-88.4
Word series, # correct	17.8	-17.1	-10.3	6
Letter series, # correct	89.7	43.1	81.8	78.4
Letter sets, # correct	87.5	45.3	84.3	75.7
PIC Self Efficacy	73.7	45.6	67.9	54.2
PIC Concern About Aging	82.8	87	78.3	53.4
SF36 Physical Health	69.9	-517.8	73.6	62.6
SF36 Mental Health	76.2	58.1	31.5	16.1

Notes: distance = Mahalonobi's squared distance, a measure of multivariate distance of each case from overall centroid using all covariates, MMSE=Mini-mental status examination, PIC=Personality In Intellectual Aging Contexts Inventory, SF36 = 36-Item Short Form RAND Health Survey, Std. Mean Diff.= Standardized Mean Difference, Var. Ratio = Variance Ratio, eCDF Mean = empirical cumulative density function mean, eCDF Max = maximum empirical cumulative density function difference, Std. Pair Dist. = standardized pair distances.

Table 3-6. Test of covariate balance under optimal pair matching, ACTIVE Reasoning and Taylor telehealth

Group:	ACTIVE controls		Taylor telehealth		
N:	105		35		
	mean	*SD*	*mean*	*SD*	*SMD*
Sex (1=male, 2 = female)	1.6	0.49	1.63	0.49	0.058
Age (years)	70.95	4.12	71.21	4.44	0.059
White (1=white, 0=not white)	0.78	0.42	0.83	0.38	0.119
Years of Education	15.43	2.73	15.49	1.77	0.025
MMSE	28.58	1.38	28.6	1.29	0.014
Vocabulary, # correct	12.54	3.98	12.74	3.37	0.054
Word series, # correct	10.41	4.78	10.49	3.54	0.018
Letter series, # correct	13.38	6.08	13.89	6.06	0.083
Letter sets, # correct	7.79	2.65	7.8	2.61	0.004
PIC Self Efficacy	32.36	3.05	32.11	3.22	0.077
PIC Concern About Aging	3.86	5.51	3.34	4.57	0.102
SF36 Physical Health	-0.3	0.89	-0.31	1.02	0.014
SF36 Mental Health	0.56	0.75	0.59	0.56	0.046

Notes: MMSE=Mini-mental status examination, PIC=Personality In Intellectual Aging Contexts Inventory, SF36 = 36-Item Short Form RAND Health Survey, *n*=sample size, *SD*=standard deviation, *SMD* = standardized mean difference (Cohen's d).

Table 3-7. Difference between the Taylor telehealth and original ACTIVE Reasoning sample.

Summary of Balance for All Data:	Means Taylor Telehealth	Means ACTIVE Reasoning	Std. Mean Diff.	Var. Ratio	eCDF Mean	eCDF Max
Distance	0.1714	0.0436	0.9464	3.419	0.3654	0.609
Sex (1=male, 2 = female)	1.6286	1.7654	-0.2791	1.3365	0.0684	0.1368
Age (years)	71.2061	73.4286	-0.5001	0.6144	0.1009	0.2331
White (1=white, 0=not white)	0.8286	0.7188	0.2913	.	0.1098	0.1098
Years of Education	15.4857	13.5083	1.116	0.4403	0.1462	0.5113
MMSE	28.6	27.2602	1.0403	0.4079	0.1675	0.3203
Vocabulary, # correct	12.7429	12.2496	0.1463	0.7384	0.0538	0.1323
Word series, # correct	10.4857	9.5188	0.2729	0.5209	0.0567	0.2301
Letter series, # correct	13.8857	10.0045	0.6402	1.1586	0.1409	0.3248
Letter sets, # correct	7.8	5.8256	0.7565	0.8725	0.1425	0.3113
PIC Self Efficacy	32.1143	30.5316	0.4912	0.7202	0.0869	0.2872
PIC Concern About Aging	3.3429	7.5609	-0.9221	0.5221	0.162	0.3353
SF36 Physical Health	-0.3082	-0.7553	0.4387	0.992	0.1467	0.2947
SF36 Mental Health	0.5862	0.3181	0.4745	0.3766	0.0818	0.218

Notes: distance = Mahalonobi's squared distance, a measure of multivariate distance of each case from overall centroid using all covariates, MMSE=Mini-mental status examination, PIC=Personality In Intellectual Aging Contexts Inventory, SF36 = 36-Item Short Form RAND Health Survey, Std. Mean Diff.= Standardized Mean Difference, Var. Ratio = Variance Ratio, eCDF Mean = empirical cumulative density function mean, eCDF Max = maximum empirical cumulative density function difference, Std. Pair Dist. = standardized pair distances.

Table 3-8. Difference between the Taylor telehealth and propensity-matched ACTIVE Reasoning group.

Summary of Balance for Matched Data:	Means Taylor Telehealth	Means ACTIVE Reasoning	Std. Mean Diff.	Var. Ratio	eCDF Mean	eCDF Max	Std. Pair Dist.
Distance	0.1714	0.1544	0.1263	1.2252	0.0078	0.1524	0.1582
Sex (1=male, 2 = female)	1.6286	1.6	0.0583	0.9919	0.0143	0.0286	0.8742
Age (years)	71.2061	70.9524	0.0571	1.1638	0.0289	0.0952	1.0045
White (1=white, 0=not white)	0.8286	0.781	0.1263	.	0.0476	0.0476	0.8339
Years of Education	15.4857	15.4286	0.0323	0.422	0.0633	0.2	1.333
MMSE	28.6	28.581	0.0148	0.8732	0.019	0.0381	0.8726
Vocabulary, # correct	12.7429	12.5429	0.0593	0.7177	0.0476	0.1143	1.1607
Word series, # correct	10.4857	10.4095	0.0215	0.5486	0.0469	0.1524	1.2312
Letter series, # correct	13.8857	13.381	0.0833	0.9957	0.0286	0.0952	0.9379
Letter sets, # correct	7.8	7.7905	0.0036	0.9686	0.017	0.0381	0.9451
PIC Self Efficacy	32.1143	32.3571	-0.0754	1.1186	0.0257	0.1143	1.0448
PIC Concern About Aging	3.3429	3.8587	-0.1128	0.69	0.0363	0.1143	1.2642
SF36 Physical Health	-0.3082	-0.2951	-0.0128	1.3162	0.039	0.1619	0.9602
SF36 Mental Health	0.5862	0.5558	0.0539	0.5679	0.0664	0.1905	1.3522

Notes: distance = Mahalonobi's squared distance, a measure of multivariate distance of each case from overall centroid using all covariates, MMSE=Mini-mental status examination, PIC=Personality In Intellectual Aging Contexts Inventory, SF36 = 36-Item Short Form RAND Health Survey, Std. Mean Diff.= Standardized Mean Difference, Var. Ratio = Variance Ratio, eCDF Mean = empirical cumulative density function mean, eCDF Max = maximum empirical cumulative density function difference, Std. Pair Dist. = standardized pair distances.

Table 3-9. Improvement in balance due to propensity matching.

Percent Balance Improvement:	Std. Mean Diff.	Var. Ratio	eCDF Mean	eCDF Max
Distance	86.7	83.5	97.9	75
Sex (1=male, 2 = female)	79.1	97.2	79.1	79.1
Age (years)	88.6	68.9	71.4	59.1
White (1=white, 0=not white)	56.6	.	56.6	56.6
Years of Education	97.1	-5.2	56.7	60.9
MMSE	98.6	84.9	88.6	88.1
Vocabulary, # correct	59.5	-9.4	11.5	13.6
Word series, # correct	92.1	8	17.4	33.8
Letter series, # correct	87	97.1	79.7	70.7
Letter sets, # correct	99.5	76.6	88.1	87.8
PIC Self Efficacy	84.7	65.8	70.4	60.2
PIC Concern About Aging	87.8	42.9	77.6	65.9
SF36 Physical Health	97.1	3322.6	73.4	45.1
SF36 Mental Health	88.6	42.1	18.9	12.6

Notes: distance = Mahalonobi's squared distance, a measure of multivariate distance of each case from overall centroid using all covariates, MMSE=Mini-mental status examination, PIC=Personality In Intellectual Aging Contexts Inventory, SF36 = 36-Item Short Form RAND Health Survey Std. Mean Diff.= Standardized Mean Difference, Var. Ratio = Variance Ratio, eCDF Mean = empirical cumulative density function mean, eCDF Max = maximum empirical cumulative density function difference, Std. Pair Dist. = standardized pair distances.

Table 3-10. Study schedule for Telehealth Inductive Reasoning Training Groups.

Battery/Training Session #	Pre-Test Battery	1	2	3	4	5	6	7	8	9	10	Post-Test Battery
Week	1-4	5	5	6	6	7	7	8	8	9	9	9-10

Table 3-11. Group number, size, and meeting days and times.

Training Groups (n = 31, after dropouts)		
Group #	Group Size	Days and Times
Group 1	n = 5	F/M at 2PM
Group 2	n = 5	T/Th at 9AM
Group 3	n = 4	W at 5:30PM & F at 11:30AM
Group 4	n = 5	Su at 11AM & W at 1PM
Group 5	n = 5	S at 1PM & M at 12PM
Group 6	n = 4	M/W at 10AM
Group 7	n = 3	M/F at 4PM

Note: M = Monday, T = Tuesday, W = Wednesday, Th = Thursday, F = Friday, S = Saturday, Su = Sunday.

Table 3-12. Session number, session content, number of practice exercises, and everyday generalization**.**

Session Number	Session Content	Number of Practice Exercises	Everyday Generalization
1	1. Discussing patterns in exercise routines and learning strategies of underlining repeats and making slash marks between pattern repetitions, 2. Studying patterns related to schedules and learning strategies of scanning, underlining, saying aloud, and drawing slashes, 3. Group exercise: understanding and remembering a prescription drug schedule, utilizing strategies (scanning and underlining), 4. End-of-session inductive reasoning assessment	3 and 1 extra letter series practice exercise after the end-of-session assessment	Finding patterns in schedules (e.g., days of the week or times of the day) and finding patterns in prescription schedules
2	1. Studying patterns related to schedules, utilizing strategies (scanning, underlining, saying aloud, and making slashes), 2. Group exercise: understanding and remembering a recycling schedule, utilizing strategies (scanning and underlining), 3. End-of-session inductive reasoning assessment	2 And 1 optional exercise in finding patterns in short schedules and 1 optional exercise in finding repeated patterns in tv channel news cycles and 1 extra letter series practice exercise after the end-of-session assessment	Finding patterns in schedules, including finding patterns in recycling schedules

Table 3-12. Continued.

Session Number	Session Content	Number of Practice Exercises	Everyday Generalization
3	1. Studying patterns related to schedules, utilizing strategies (scanning, underlining, reading aloud, and making slashes), 2. Applying these strategies (scanning, underlining, reading aloud, and making slashes) to schedules with abbreviations (e.g., “Mo” for Monday), 3. Group exercise: understanding and remembering a medication schedule, utilizing strategies (scanning and underlining), 4. End-of-session inductive reasoning assessment	3 and 1 optional exercise of finding pattens in schedules	Finding patterns in schedules, including medication schedules
4	1. Psychoeducation on the importance of letter patterns for ordering from mail-order catalogues, understanding model or serial numbers, and knowing radio and T.V. channels, 2. Determining if a name comes before or after a listing, 3. Practice finding the pattern in a series of letters, utilizing strategies (scanning, underlining, and saying aloud), 4. Completing an exercise in following directions in daily activities (taking a laxative) by applying strategies (scanning, underlining, and reading aloud) to directions and questions (e.g., “can apple cider be mixed with the laxative”), 5. End-of-session inductive reasoning assessment	3 and 1 extra letter series practice exercise after the end-of-session assessment	Understanding model or serial numbers and radio and T.V. Channels, Looking up names in phone books, Finding patterns in instructions for taking a laxative
5	1. Psychoeducation on patterns in daily life (songs, knitting, tv schedules), 2. Practice finding the pattern in a series of letters, 3. Identifying skip patterns in schedules and letter series (using tick marks to indicate skips), 4. Group exercise: understanding medicine labels, utilizing strategies (scanning and underlining techniques), 5. End-of-session inductive reasoning assessment	4 and 1 extra letter series practice exercise after the end-of-session assessment	Finding patterns in songs, knitting, and T.V. schedules; identifying patterns, including skips, in schedules; understanding medicine labels

Table 3-12. Continued.

Session Number	Session Content	Number of Practice Exercises	Everyday Generalization
6	1. Practice using strategies (scanning, underlining, saying aloud, making tick marks, and making slashes) to identify the next item in a series of letters, 2. Group exercise: finding health and government service information in tables, while utilizing strategies (scanning table and underlining key words) 3. Using strategies (saying aloud, underlining, and drawing slashes) to complete schedules, 4. End-of-session inductive reasoning assessment	4 and 1 extra letter series practice exercise after the end-of-session assessment	Finding information in tables (health, finances, travel, and nutrition); understanding patterns in schedules.
7	1. Psychoeducation on how practice improves basic mental abilities and how practice on exercise similar to everyday tasks (medication and recycling charts and understanding medicine bottle labels) helps apply basic training to things participants do everyday, 2. Practice using strategies (scanning, underlining, saying aloud, making tick marks, and making slashes) to identify the next item in a series of letters, 3. Group exercise: applying strategies (scanning and underlining) to nutrition facts charts, 4. Using strategies (saying aloud, underlining, and drawing slashes) to complete schedules, 5. End-of-session inductive reasoning assessment	3 and 1 optional finding the pattern in schedules, coin sets, telephone numbers, names, and zip codes and 1 extra letter series practice exercise after the end-of-session assessment	Understanding nutrition facts charts, understanding patterns in schedules

Table 3-12. Continued.

Session Number	Session Content	Number of Practice Exercises	Everyday Generalization
8	1. Practice using strategies (scanning, underlining, saying aloud, making tick marks, and making slashes) to identify the next item in a series of letters, 2. Group exercise: Studying and learning how to understand the organization of a medication and a transportation chart, utilizing strategies (scanning, underlining, and reading aloud), 3. Finding the letter series set that is different, utilizing strategies (scanning, underlining, and reading aloud), 4. End-of-session inductive reasoning assessment	3 and 1 optional filling in the missing order numbers in a table exercise and 1 optional finding what comes next in the schedule and 1 extra letter series practice exercise after the end-of-session assessment	Understanding medication and transportation charts.
9	1. Finding the letter series set that is different, utilizing strategies (scanning, underlining, reading aloud, and tick marks), 2. Group exercise: Studying and learning how to read and understand a nutrition chart and bus schedule, utilizing strategies (scanning, underlining, and reading aloud), 3. Practice using strategies (scanning, underlining, saying aloud, making tick marks, and making slashes) to identify the next item in a series of letters, 4. End-of-session inductive reasoning assessment	3 and 1 optional finding patterns in television news channel schedules and 1 extra letter series practice exercise after the end-of-session assessment	Understanding nutrition charts and bus schedules

Table 3-12. Continued.

Session Number	Session Content	Number of Practice Exercises	Everyday Generalization
10	1. Psychoeducation regarding the benefits of continued mental exercise in daily life, 2. Finding the letter series set that is different, utilizing strategies (scanning, underlining, reading aloud, and using tick marks), 3. Using strategies (scanning, underlining, and drawing slashes) to complete schedules, 4. Group exercise: using strategies (scanning, underlining, and reading aloud) to identify the pattens in a schedule of weekly senior center activities or senior van services, 5. Practice finding the next letter in a series utilizing strategies (scanning, underlining, and drawing slashes), 6. End-of-session inductive reasoning assessment	4 and 1 extra letter series practice exercise after the end-of-session assessment	Understanding patterns in schedules and weekly senior activity schedules and senior van services.

Table 3-13. Measures.

Method of Measurement in Present Study	*Targeted Domain*	*Purpose or Transfer Category*	*Occasion*
Consent Form	Consent	--	Pre-test Only
Demographic/Personal Data Form	Demographics and Personal Data	Covariates and Inclusion	Pre-test Only
Telehealth Readiness Form	Telehealth Readiness	Inclusion	Pre-test Only
Mini-Mental State Exam (Folstein et al., 1975)	General Cognition	Inclusion	Pre-test Only
Short-Form Health Survey; SF36GH (Ware Jr & Sherbourne, 1992)	Self-Rated Health	Covariate	Pre-Test Only
Training Summary Data Form	Inductive Reasoning Process Scores, Environmental Difficulty, Trainer Perceptions of Improvement, Scheduled Date, Actual Date, Actual Time, and Exposure Duration.	Covariate	During and After Each Session
Word Series (Gonda & Abilities, 1985) Letter Series (Thurstone & Thurstone, 1947) Letter Sets (Ekstrom et al., 1976)	Inductive Reasoning	Near Transfer	Pre-test and Post-test
Vocabulary (Ekstrom et al., 1976)	Language	Far transfer; no Transfer Expected	Pre-test and Post-test
Personality in Intellectual-Aging Contexts (Lachman et al., 1982)	Self-efficacy over and concern about intellectual aging	Far transfer; no Transfer Expected	Pre-test and Post-test
Modified Telehealth Usability Questionnaire	Telehealth Usability and Feasibility	Outcome Measure	Post-test Only

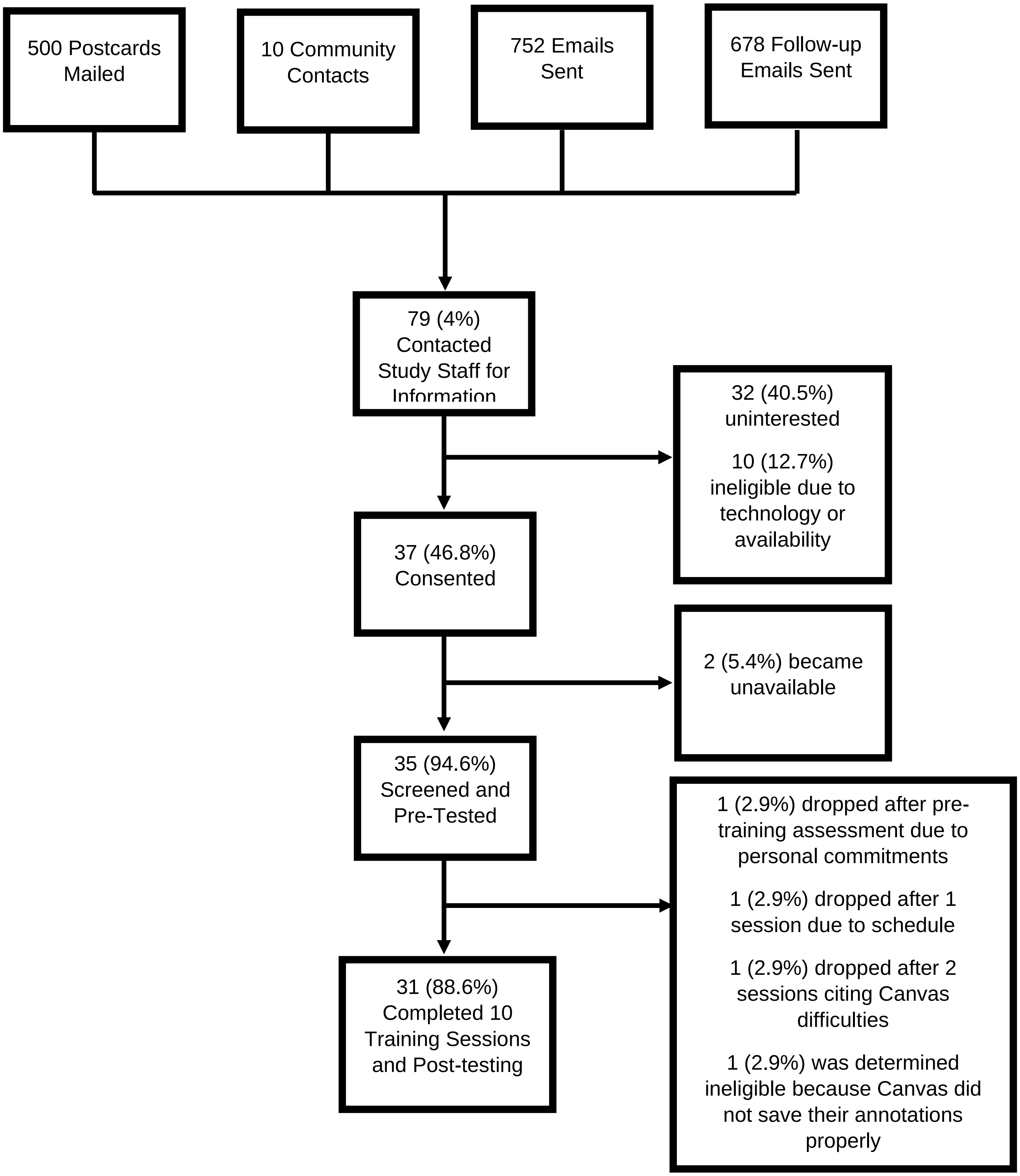


Figure 3-1. Consort diagram describing the flow of participants from recruitment through post-testing.

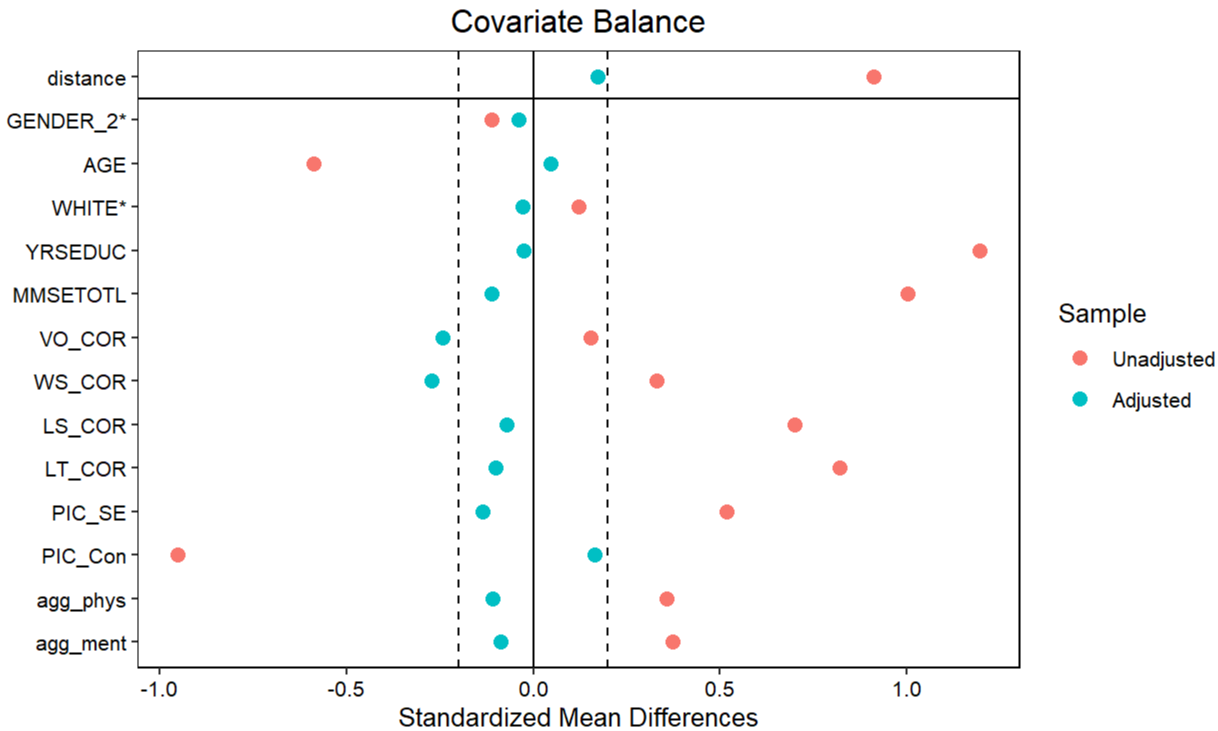


Figure 3-2. Standardized Mean Differences on each Variable between the Total ACTIVE Control Sample and the Taylor Telehealth Sample (Pink) and for the Propensity-Matched ACTIVE Control Sample and Taylor Telehealth (Blue). Notes: GENDER_2 = Sex (1=Male, 2=Female), WHITE = White race (1=yes, 0 = no), YRSEDUC = Years of Education, MMSETOTL = Mini Mental Status Examination, VO_COR = Vocabulary, # correct, WS_COR = Word Series, # correct, LS_COR = Letter Series, # correct, LT_COR = Letter Sets, # Correct, PIC_SE = Personality in Intellectual Aging Contexts Self Efficacy; PIC_Con = Personality in Intellectual Aging Contexts Concern about Aging, agg_phys = Aggregate physical health, 36-item Short Form RAND Health Survey, agg_ment = Aggregate mental health, 36-item Short Form RAND Health Survey.

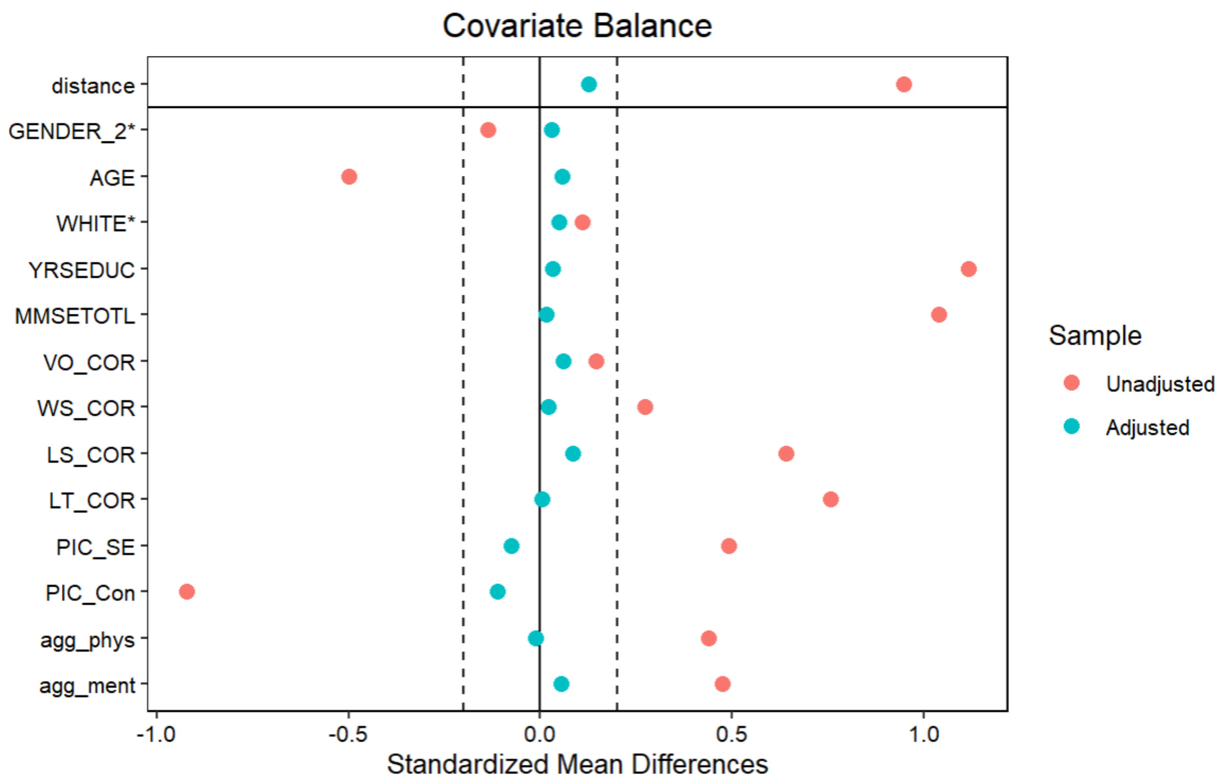


Figure 3-3. Standardized Mean Differences on each Variable between the Total ACTIVE Reasoning Sample and the Taylor Telehealth Sample (Pink) and for the Propensity-Matched ACTIVE Control Sample and Taylor Telehealth (Blue). Notes: GENDER_2 = Sex (1=Male, 2=Female), WHITE = White race (1=yes, 0 = no), YRSEDUC = Years of Education, MMSETOTL = Mini Mental Status Examination, VO_COR = Vocabulary, # correct, WS_COR = Word Series, # correct, LS_COR = Letter Series, # correct, LT_COR = Letter Sets, # Correct, PIC_SE = Personality in Intellectual Aging Contexts Self Efficacy; PIC_Con = Personality in Intellectual Aging Contexts Concern about Aging, agg_phys = Aggregate physical health, 36-item Short Form RAND Health Survey, agg_ment = Aggregate mental health, 36-item Short Form RAND Health Survey.

Monday
Monday
Tuesday
Wednesday
Wednesday
Thursday
Friday
Friday
Saturday

Figure 3-4. Example underlining word repeats and drawing slashes between pattern repeats strategies.

Time Schedules

The time schedules in our daily lives follow a pattern. Below is a schedule for a Cable TV News Channel. The same schedule is repeated every 30 minutes throughout the day and evening. Describe the pattern. If you remember the schedule, you will know when to turn to the channel for sports or for weather any day of the week or any hour of the day.

00:00	NEWS
00:10	WEATHER
00:15	SPORTS
00:25	ENTERTAINMENT
00:30	NEWS
00:40	WEATHER
00:45	SPORTS
00:55	ENTERTAINMENT
00:00	NEWS

Figure 3-5. Example television schedule demonstrating its pattern. Participants are asked questions about this pattern.

Figure 3-6. Example of how the participant can use the paintbrush feature in Canvas to underline repeated letters or words, make slashes between groups of letters or words, and circle answer choices.

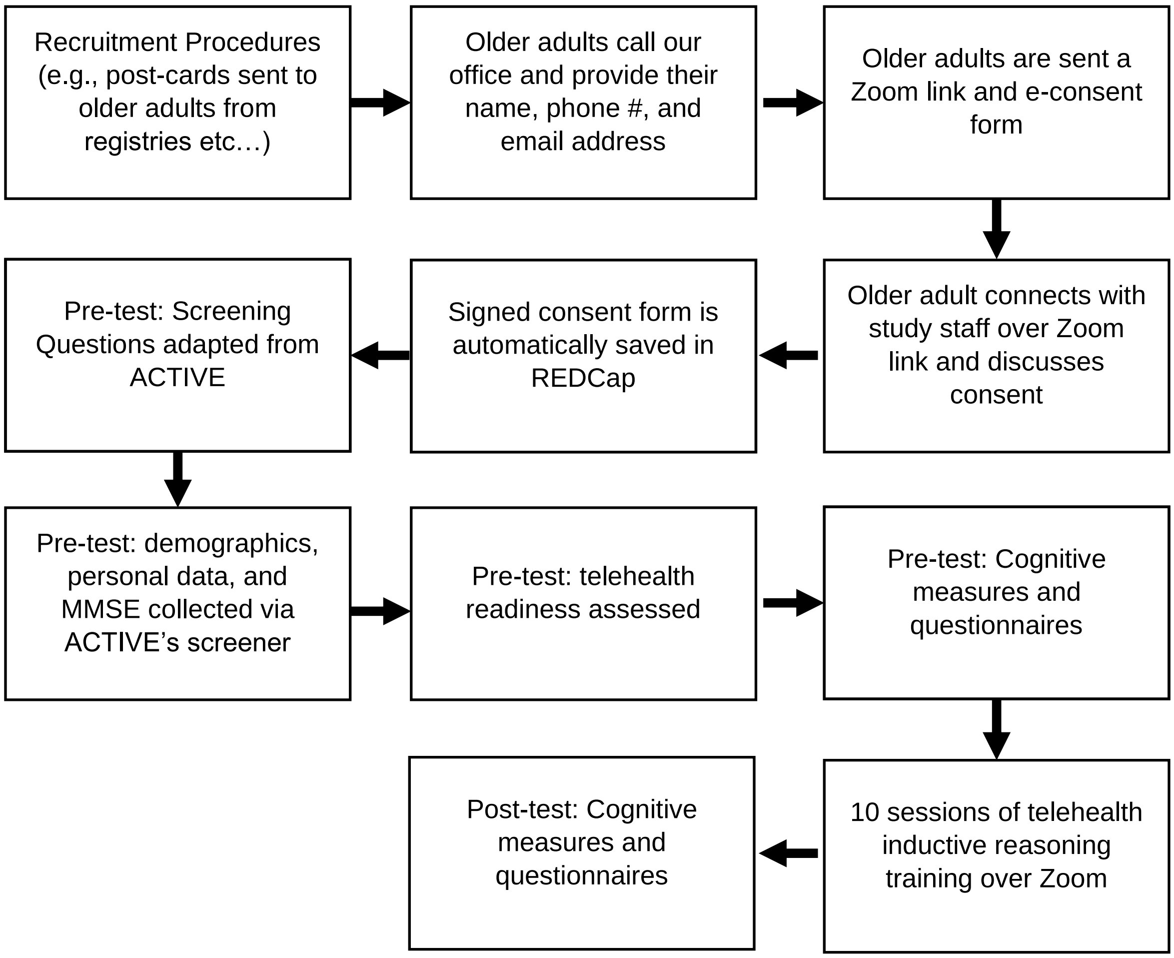


Figure 3-7. Procedural Sequence Flow Diagram.

CHAPTER 4
RESULTS

Overview

Below, the goal of Aim 1 was to provide a summary of perceptions of and compliance with a telehealth-delivered cognitive training intervention in inductive reasoning. We also provided a descriptive summary of findings regarding telehealth readiness. Aim 2 was to determine if a telehealth cognitive training intervention in inductive reasoning was effective in improving inductive reasoning, relative to propensity-matched untrained, no-contact controls. Aim 3 was to conduct a descriptive comparison to examine whether a telehealth-delivered cognitive training intervention in inductive reasoning was comparable to propensity-matched participants from traditional face-to-face inductive reasoning training in ACTIVE. As each aim was investigated, follow-up analyses were identified to try to shed light on some unexpected patterns. These are presented as additional follow-up analyses under each aim. Demographic information can be found in the Methods section.

Aim 1

Summary of Telehealth Readiness

Prior to investigating the main question of the aim (perception of and compliance with the intervention), we first briefly looked at telehealth readiness based on our screening inventory. Telehealth readiness was assessed at pretest via a question asking whether they had used Zoom before with family, friends, or healthcare providers, questions assessing their ability to perform basic Zoom features, and a question assessing their ability to annotate assignments in Canvas.

As shown in Figure 4-1, on the question asking whether the participants had used Zoom with family, friends, or healthcare providers, 25 of the 31 (80.6%) reported having used Zoom before and 6 of the 31 (19.4%) reported never using Zoom before. All 4 non-compliant participants (those who did not complete pretest, all 10 training sessions, and posttest) individuals had used Zoom before.

There were 6 questions assessing the participants' ability to use Zoom: 1. Ability to share screen, 2. Ability to switch between gallery and speaker views, 3. ability to mute and unmute, 4. ability to turn on and off video, 5. ability to use chat, and 6. ability to optimize volume. Participants were rated whether they could not perform the ability, they could with assistance, or they could with no assistance needed. These variables were summed to compute an overall Zoom familiarity score. 15 (48.4%) of the 31 participants performed perfectly (no assistance needed on all of the 6 abilities tested), 4 participants (12.9%) required assistance on only one of the 6 abilities, 6 (19.4%) required assistance on two of the 6 abilities, 3 (9.77%) required assistance on three of the 6 abilities, 1 (3.2%) required assistance on four of the 6 abilities, 1 (3.2%) required assistance on 5 five of the 6 abilities, and 1 (3.2%) was unable to perform any of the abilities. Results for the individual questions and the overall Zoom familiarity can be found in Figures 4-2 and 4-3. All 4 non-compliant participants performed perfectly (not requiring any assistance on all the 6 abilities tested).

On a question assessing the participants' ability to annotate assignments in Canvas, 12 (38.7%) annotated without assistance and 19 (61.3%) annotated with assistance (see Figure 4-4). 2 of the non-compliant participants (50%) annotated without assistance and the other 2 (50%) with assistance.

Participant Perceptions of a Telehealth-Delivered Cognitive Training Intervention in Inductive Reasoning

The participants' perceptions of the telehealth-delivered cognitive training intervention in inductive reasoning were measured at posttest via a Telehealth Usability Questionnaire, with 20 questions, assessing 6 domains: 1. Usefulness, 2. Ease of Use & Learnability, 3. Interface Quality, 4. Interaction Quality, 5. Reliability, and 6. Satisfaction & Future Use. Answer choices for all 20 questions were the same, a Likert scale ranging from Strongly Disagree, Disagree, Neutral, Agree, and Strongly Agree. These results are displayed as stacked horizontal bar charts in Figures 4-5 through 4-10. Results from the 31 compliant participants are shown. The decision to not include the 4 non-compliant participants was because the researchers felt that those participants were not capable of answering the questions regarding their sentiments on the entire intervention experience. Within the Usefulness domain (Figure 4-5), most participants strongly agreed with the questions; however, for the question asking participants if telehealth improved their access to health care services, a higher proportion answered "neutral", likely reflecting the fact that this cognitive training intervention was not seen as related to other healthcare services. For the Ease of Use & Learnability domain (Figure 4-6), most participants strongly agreed with the questions that it was easy to learn and simple to use Zoom, and over 90% strongly agreed or simply agreed. For the Interface Quality domain (Figure 4-7), roughly half strongly agreed that the Zoom interface is simple and easy to understand that they enjoyed using the Zoom interface, and over 85% strongly agreed or agreed. For the Interaction Quality domain (Figure 4-8), over 70% strongly agreed that they were able to express themselves effectively, could hear the trainer, and could easily talk to the trainer, and over 90% strongly agreed or agreed with those statements. However, for the question asking if this system is able to do everything they would want it to be able to do, slightly fewer than half of the participants strongly agreed, 16% agreed, 23% were neutral, and 13% disagreed. For the Reliability domain (Figure 4-9), over 75% strongly agreed that when using the telehealth system, they can see the trainer as well as if they met in person, and over 90% strongly agreed or agreed with that statement. However, for the question asking if they think the visits provided over the telehealth system are the same as in-person visits, only 19% strongly agreed, 29% agreed, 16% were neutral, 29% disagreed, and 6% strongly disagreed. For the Satisfaction & Future Use domain (Figure 4-10), more than 50% strongly agreed with all 6 statements, and over 90% strongly agreed or agreed with all 6 statements.

Participants Compliance with a Telehealth-Delivered Cognitive Training Intervention in Inductive Reasoning

The telehealth inductive reasoning study had 31 participants who attended all or most training sessions and the posttest, representing 88.6% of the initially enrolled sample of N = 35. Thus, there were 4 non-compliant participants. 1 completed only the pretest and was removed from further participation due to technological limitations. 1 completed the pretest and 2 training sessions before dropping out and not doing the posttest. 1 completed the pretest and posttest, but none of the training sessions. And 1 other completed the pretest, 1 training session, and the posttest. These individuals are also shown in the Consort Diagram (Figure 3-1).

In the N = 31 sample of those who actually received Telehealth Reasoning training, with compliance being defined as having completed over 80% of the 10 training sessions, all 31 of those participants (100%) completed all 10 training sessions. However, as noted below, extensive (possibly excessive) accommodations were made to participants (e.g., evening and weekend sessions, makeups, etc) to achieve this perfect compliance.

Aim 2

The means and standard deviations of measures of Word Series, Letter Series, Letter Sets, Vocabulary, Self-Efficacy, and Concern about Aging can be found in Table 4-1.

Telehealth Reasoning Versus Propensity-Matched No-Contact Controls

As described in Table 4-2, an ANOVA revealed a non-significant interaction between the effects of Occasion (pretest and posttest) and groups (Taylor telehealth trained and ACTIVE controls) on Word Series Score [$F(1, 116) = 1.145$, $p = 0.287$, *partial eta-squared* = 0.010; Taylor telehealth trained pretest mean = 10.90 and posttest mean = 12.87; ACTIVE controls pretest mean = 11.49 and posttest mean = 12.87]. This can be seen in Figure 4-11. Also described in Table 4-2, simple main effects analysis revealed group (Taylor telehealth trained and ACTIVE controls) did not have a significant effect on Word Series score [$F(1, 116) = 0.912$, $p = 0.342$, *partial eta-squared* = 0.008; Taylor telehealth mean Word Series score across occasions = 8.69; ACTIVE controls mean Word Series score across occasions = 8.36]. Further described in Table 4-2, simple main effects analysis showed that occasion (pretest to posttest) did have a significant effect on Word Series score [$F(1, 116) = 56.529$, $p < 0.001$, *partial eta-*

squared = 0.328]. The participants improved an average of 2.294 points from pretest to posttest (pretest mean = 11.199, posttest mean = 13.493).

However, as described in Table 4-2, an ANOVA revealed a significant interaction between the effects of Occasion (pretest and posttest) and groups (Taylor telehealth trained and ACTIVE controls) on Letter Series Score [$F(1, 116) = 9.81$, $p = 0.002$, *partial eta-squared* = 0.08; Taylor telehealth trained pretest mean = 14.06 and posttest mean = 17.65; ACTIVE controls pretest mean = 14.40 and posttest mean = 15.78]. This interaction revealed that persons in the Taylor telehealth trained group evinced more improvement on Letter Series from pretest to posttest than ACTIVE no-contact controls. This can be seen in Figure 4-12. Simple main effects analysis revealed group (Taylor telehealth trained and ACTIVE controls) did not have a significant effect on Letter Series score [$F(1, 116) = 0.45$, $p = 0.504$, *partial eta-squared* = 0.004; Taylor telehealth mean Letter Series score across occasions = 15.85; ACTIVE controls mean Letter Series score across occasions = 15.09]. Simple main effects analysis showed that occasion (pretest to posttest) did have a significant effect on Letter Series score [$F(1, 116) = 49.75$, $p < 0.001$, *partial eta-squared* = 0.30]. The participants improved an average of 2.480 points from pretest to posttest (pretest mean = 14.233, posttest mean = 16.713).

As described in Table 4-2, an ANOVA revealed a significant interaction between the effects of Occasion (pretest and posttest) and groups (Taylor telehealth trained and ACTIVE controls) on Letter Sets Score [$F(1, 116) = 8.69$, $p = 0.004$, *partial eta-squared* = 0.07; Taylor telehealth trained pretest mean = 7.77 and posttest mean = 9.35; ACTIVE controls pretest mean = 8.25 and posttest mean = 8.47]. This interaction revealed that persons in the Taylor telehealth trained group evinced more improvement on Letter Sets from pretest to posttest than ACTIVE no-contact controls. This can be seen in Figure 4-13. Simple main effects analysis revealed group (Taylor telehealth trained and ACTIVE controls) did not have a significant effect on Letter Sets score [$F(1, 116) = 0.14$, $p = 0.707$, *partial eta-squared* = 0.001; Taylor telehealth mean Letter Sets score across occasions = 8.57; ACTIVE controls mean Letter Sets score across occasions = 8.36]. Simple main effects analysis showed that occasion (pretest to posttest) did have a significant effect on Letter Sets score [$F(1, 116) = 15.15$, $p < 0.001$, *partial eta-squared* = 0.12]. The participants improved an average of 0.900 points from pretest to posttest (pretest mean = 8.014, posttest mean = 8.913).

Personality in intellectual aging Contexts

As described in Table 4-3, an ANOVA revealed a significant interaction between the effects of Occasion (pretest and posttest/1-year follow-up) and groups (Taylor telehealth trained and ACTIVE controls) on Self Efficacy score [$F(1, 104) = 11.169$, $p = 0.001$, *partial eta-squared* = 0.097], with Taylor telehealth participants self-reporting 1.708 more points of improvement in self-efficacy from pretest to posttest compared to ACTIVE controls (Taylor telehealth trained pretest mean = 32.11 and posttest mean = 33.58; ACTIVE controls pretest mean = 32.48 and 1-year follow-up mean = 32.24). This can be seen in Figure 4-14. Also described in Table 4-3, simple main effects analysis revealed group (Taylor telehealth trained and ACTIVE controls) did not have a significant effect on Self Efficacy score [$F(1, 104) = 0.624$, $p = 0.432$, *partial eta-squared* = 0.006; Taylor telehealth trained mean Self Efficacy score across pretest and posttest = 32.85; ACTIVE controls mean Self Efficacy score across pretest and 1-year follow-up = 32.36]. Simple main effects analysis showed that occasion (pretest to posttest/1-year follow-up) did have a significant effect on Self Efficacy score [$F(1, 104) = 5.773$, $p = 0.018$, *partial eta-squared* = 0.053]. The participants improved an average of 0.614 points from pretest to posttest/1-year follow-up (pretest mean across groups = 32.30; posttest/1-year follow-up mean across groups = 32.91).

As described in Table 4-3, an ANOVA revealed a significant interaction between the effects of Occasion (pretest and posttest/1-year follow-up) and groups (Taylor telehealth trained and ACTIVE controls) on Concern About Aging Score [$F(1, 104) = 6.303$, $p = 0.014$, *partial eta-squared* = 0.057]. The Taylor telehealth trained participants reported 0.96 points less concern about aging at posttest/1-year follow-up compared to pretest, while the ACTIVE controls reported .825 points more concern about aging at the 1-year follow-up compared to pretest (Taylor telehealth trained pretest mean = 3.234 and posttest mean = 2.274; ACTIVE controls pretest mean = 2.557 and 1-year follow-up mean = 3.382). This can be seen in Figure 4-15. Also described in Table 4-3, simple main effects analysis revealed group (Taylor telehealth trained and ACTIVE controls) did not have a significant effect on Concern About Aging Score [$F(1, 104) = 0.045$, $p = 0.833$, *partial eta-squared* < 0.001; Taylor telehealth trained mean concern about aging across pretest and posttest = 2.754; ACTIVE controls mean concern about aging across pretest and 1-year follow-up = 2.969]. Simple main effects analysis showed that occasion (pretest to posttest/1-year follow-up) did have a significant effect on Concern about Aging score [$F(1, 104) = 0.036$, $p = 0.851$, *partial eta-squared* < 0.001; mean concern about aging at pretest across the groups = 2.895; mean concern about aging at posttest/1-year follow-up across the groups = 2.828].

Vocabulary

As described in Table 4-3, an ANOVA revealed a non-significant interaction between the effects of Occasion (pretest and posttest/1-year follow-up) and groups (Taylor telehealth trained and ACTIVE controls) on Vocabulary Score [$F(1, 104) = 0.679$, $p = 0.412$, *partial eta-squared* = 0.006; Taylor telehealth trained mean pretest Vocabulary score = 12.807 and mean posttest Vocabulary score= 13.520; ACTIVE controls mean pretest Vocabulary score = 13.573 and mean 1-year follow-up Vocabulary score = 13.950]. This can be seen in Figure 4-16. Also described in Table 4-3, simple main effects analysis revealed group (Taylor telehealth trained and ACTIVE controls) did not have a significant effect on Vocabulary score [$F(1, 104) = 0.830$, $p = 0.364$, *partial eta-squared* = 0.008; Taylor telehealth trained mean Vocabulary score across pretest and posttest = 13.161; ACTIVE Control mean Vocabulary score across pretest and 1-year follow-up = 13.760]. Further described in Table 4-3, simple main effects analysis showed that occasion (pretest to posttest) did have a significant effect on Vocabulary score [$F(1, 104) = 7.042$, $p = 0.009$, *partial eta-squared* = 0.063]. The participants improved an average of 0.542 points from pretest to posttest/1-year follow-up (Pretest mean Vocabulary score across groups = 13.190; Posttest/1-year follow-up mean Vocabulary score across groups = 13.731).

Additional Follow-up Analysis: Number of Attempts

The means and standard deviations of the number of attempts on Word Series, Letter Series, Letter Sets, Vocabulary, Self-Efficacy, and Concern about Aging can be found in Table 4-4.

To better understand whether the two groups (Taylor telehealth trained vs. ACTIVE no-contact controls) differed to the degree that they optimized speed, the number of attempted inductive reasoning items at pretest and posttest were compared (see Table 4-5). This comparison between the two groups on the number of attempted problems is important because, for example, if the improvement in the number of correct responses is perfectly parallel to the number of problems attempted, it suggests that training improved the participants speed of responding or efficiency. However, if the number of correct responses increased but the number of attempts was the same at pretest and posttest, that would imply that the intervention made them more accurate.

As depicted in Table 4-5, an ANOVA revealed a significant interaction between the effects of Occasion (pretest and posttest) and groups (Taylor telehealth trained and ACTIVE no-contact controls) on the number of attempted items on Word Series [$F(1, 116) = 6.643$, $p = 0.011$, *partial eta-squared* = 0.054], revealing that the persons in the ACTIVE control group evinced a slightly greater increase in attempts from pretest to posttest than Taylor telehealth trained participants. This can be seen in Figure 4-17. As depicted in Table 4-5, an ANOVA revealed non-significant interaction between the effects of Occasion (pretest and posttest) and groups (Taylor telehealth trained and ACTIVE no-contact controls) on the number of attempted items on Letter Series [$F(1, 116) = 0.139$, $p = 0.710$, *partial eta-squared* = 0.001]. Instead, there was a main effect of Occasion ($F(1,116) = 27.799$, $p < .001$, partial eta-squared = 0.193, reflecting that both groups increased in attempts from pretest to posttest. This can be seen in Figure 4-18. As depicted in Table 4-5, an ANOVA revealed a significant interaction between the effects of Occasion (pretest and posttest) and groups (Taylor telehealth trained and

ACTIVE no-contact controls) on the number of attempted items on Letter Sets [F(1, 116) = 4.619, p = 0.034, *partial eta-squared* = 0.038], revealing that, opposite to what was observed for Word Series, the persons in the Taylor telehealth trained group increased more in attempts from pretest to posttest than no-contact ACTIVE controls. This can be seen in Figure 4-19.

For the most part, group differences in changes in attempts did not help to explain group differences in changes in number of correct items. Only for Letter Sets were these changes congruent (i.e., telehealth-trained participants experienced greater improvement in attempts and rights). For Letter Series, despite the greater improvement of telehealth trained participants in number of correct items, there was not a congruent disproportionate improvement in attempts; both groups improved equally. Most difficult to understand was Word Series, where there was not group difference in improvement in number of correct items but the untrained ACTIVE controls actually improvement more in number of attempts. We have reason to believe that mode effects (i.e., the online administration in telehealth-trained persons) may have particularly affected Word Series, which help may account for the unexpected group difference in changes in number of attempts. The discussion section will consider why the Word Series measure, in particular, may have been less likely to increase in attempts using the telehealth/Canvas platform.

Aim 3

The means and standard deviations of measures of Word Series, Letter Series, Letter Sets, Vocabulary, Self-Efficacy, and Concern about Aging can be found in Table 4-1.

Telehealth Reasoning Versus Propensity-Matched ACTIVE Inductive Reasoning Trained

Both the telehealth trained participants and the propensity-matched ACTIVE in-person trained participants completed 100% of the training sessions; thus, they had equivalent dosages of training. As described in Table 4-6, an ANOVA revealed a significant interaction between the effects of Occasion (pretest and posttest) and groups (Taylor telehealth trained and ACTIVE in-person trained) on Word Series Score [F(1, 116) = 25.681, p = < 0.001, *partial eta-squared* = 0.181]. This interaction revealed that persons in the ACTIVE in-person trained group evinced more improvement on Word Series from pretest to posttest than the Taylor telehealth-trained participants (Taylor telehealth trained pretest mean = 10.90 and posttest mean = 12.87; ACTIVE in-person trained pretest mean = 10.47 and posttest mean = 15.99). This can be seen in Figure 4-20.

As described in Table 4-6, an ANOVA revealed a non-significant interaction between the effects of Occasion (pretest and posttest) and groups (Taylor telehealth trained and ACTIVE in-person trained) on Letter Series Score [F(1, 116) = 1.24, p = 0.226, *partial eta-squared* = 0.01; Taylor telehealth trained pretest mean = 14.06 and posttest mean = 17.65; ACTIVE in-person trained pretest mean = 13.49 and posttest mean = 17.95]. This can be seen in Figure 4-21. Although the interaction was not significant, the main effect of occasion was [$F(1, 116) = 104.4$, $p < 0.001$, *partial eta-squared* = 0.47; overall pretest mean = 13.78 and overall posttest mean = 17.80]. The effect of Group (Taylor telehealth trained vs ACTIVE in-person trained) was not significant [$F(1, 116) = 0.01$, $p = 0.910$, *partial eta-squared* < 0.001; Taylor telehealth trained mean across occasions = 15.86 and; ACTIVE in-person trained mean across occasions = 15.72].

As described in Table 4-6, an ANOVA revealed a non-significant interaction between the effects of Occasion (pretest and posttest) and groups (Taylor telehealth trained and ACTIVE in-person trained) on Letter Sets Score [$F(1, 116) = 1.29$, $p = 0.259$, *partial eta-squared* = 0.01; Taylor telehealth trained pretest mean = 7.77 and posttest mean = 9.35; ACTIVE in-person trained pretest mean = 7.83 and posttest mean = 8.80]. This can be seen in Figure 4-22. Although the interaction was not significant, the main effect of occasion was [$F(1, 116) = 23.14$, $p < 0.001$, *partial eta-squared* = 0.17; overall pretest mean = 7.80 and overall posttest mean = 9.08]. The effect of Group (Taylor telehealth trained vs ACTIVE in-person trained) was not significant [$F(1, 116) = 0.22$, $p = 0.638$, *partial eta-squared* = 0.002; Taylor telehealth trained mean across occasions = 8.57 and; ACTIVE in-person trained mean across occasions = 8.32].

Personality in intellectual aging Contexts

As described in Table 4-7, an ANOVA revealed a significant interaction between the effects of Occasion (pretest and posttest or 1-year follow-up, in the case of the ACTIVE controls) and groups (Taylor telehealth trained and ACTIVE in-person) on self-rated self-efficacy for cognitive aging [$F(1, 107) = 5.896$, $p = 0.017$, *partial eta-squared* =

0.052; Taylor telehealth trained pretest mean = 32.11 and posttest mean = 33.58; ACTIVE in-person trained pretest mean = 32.51 and 1-year follow-up mean = 32.56], revealing that the persons in the Taylor telehealth trained group evinced more self-rated improvement in self-efficacy than ACTIVE's in-person trained from pretest to posttest (or 1 year follow-up, in the case of the ACTIVE controls). This can be seen in Figure 4-23.

As described in Table 4-7, an ANOVA revealed a non-significant interaction between the effects of Occasion (pretest and posttest or 1-year follow-up, in the case of the ACTIVE in-person trained) and groups (Taylor telehealth trained and ACTIVE in-person) on self-rated concern about aging [$F(1, 107) = 0.765$, $p = 0.384$, *partial eta-squared* = 0.007; Taylor telehealth trained pretest mean = 3.23 and posttest mean = 2.27; ACTIVE in-person trained pretest mean = 3.57 and 1-year follow-up mean = 3.23]. This can be seen in Figure 4-24. The main effect of occasion was also not significant [$F(1, 107) = 3.45$, $p = 0.066$, *partial eta-squared* = 0.031; overall pretest mean = 3.40 and overall posttest mean = 2.75]. The effect of Group (Taylor telehealth trained vs ACTIVE in-person trained) was also not significant [$F(1, 107) = 0.346$, $p = 0.558$, *partial eta-squared* = 0.003; Taylor telehealth trained mean across occasions = 2.75 and; ACTIVE in-person trained mean across occasions = 3.40].

Vocabulary

As described in Table 4-7, an ANOVA revealed a non-significant interaction between the effects of Groups (Taylor telehealth trained and ACTIVE in-person) and Occasion (pretest and posttest or 1-year follow-up, in the case of the ACTIVE in-person trained group) on Vocabulary Score [$F(1, 107) = 0.222$, $p = 0.638$, *partial eta-squared* =

0.002; Taylor telehealth trained mean Vocabulary score at pretest = 12.81 and posttest = 13.52; ACTIVE in-person trained at pretest = 12.76 and 1-year follow-up = 13.69]. This can be seen in Figure 4-25. Although the interaction was not significant, the main effect of occasion (pretest to posttest) was [$F(1, 107) = 11.749$, $p < 0.001$, *partial eta-squared* = 0.099], showing that both groups experienced significant improvement. The participants improved an average of 0.82 points from pretest to posttest/1-year follow-up (pretest mean across groups = 12.78; posttest/1-year follow-up mean across groups = 13.60). However, average performance across occasions did not differ significantly by group [$F(1, 107) = 0.008$, $p = 0.929$, *partial eta-squared* < .0001; Taylor telehealth trained mean Vocabulary score across pretest and posttest = 13.16; ACTIVE in-person trained mean Vocabulary score at pretest and 1-year follow-up = 13.22].

Additional Follow-up Analyses

Number of attempts

The means and standard deviations of the number of attempts on Word Series, Letter Series, Letter Sets, Vocabulary, Self-Efficacy, and Concern about Aging can be found in Table 4-4.

As with Aim 2, the comparison of Taylor telehealth trained to ACTIVE no-contact controls on the number of attempts, we repeated this analysis in Aim 3. As depicted in Table 4-8, an ANOVA revealed a significant interaction between the effects of Occasion (pretest and posttest) and groups (Taylor telehealth trained and ACTIVE in-person trained) on the number of attempted items on Word Series [$F(1, 116) = 18.088$, $p < 0.001$ *partial eta-squared* = 0.135], revealing that the persons in the Taylor telehealth trained group evinced slightly more attempts from pretest (13.29) to posttest (13.90), while the ACTIVE in-person increased much more in number of attempts from pretest (14.41) to posttest (18.74). This can be seen in Figure 4-26.

As depicted in Table 4-8, there was not a significant interaction between the effects of Occasion (pretest and posttest) and groups (Taylor telehealth trained and ACTIVE in-person trained) on the number of attempted items on Letter Series [$F(1, 116) = 0.075$, $p = 0.785$ *partial eta-squared* = 0.001]. There was however a significant main effect of Occasion, such that both groups improved; $F(1,116) = 27.804$, $p < .001$ partial eta squared = 0.193). The Taylor telehealth trained group improved from pretest (16.97) to posttest (18.97), as did the ACTIVE in-person trained group, increasing from pretest (18.65) to posttest (20.86). This can be seen in Figure 4-27.

As depicted in Table 4-8, there was also not a significant interaction between the effects of Occasion (pretest and posttest) and groups (Taylor telehealth trained and ACTIVE in-person trained) on the number of attempted items on Letter Sets [$F(1, 116) = 1.330$, $p < 0.251$ *partial eta-squared* = 0.011]. There was again a significant main effect of Occasion, such that both groups improved; $F(1,116) = 18.51$, $p < .001$ partial eta squared = 0.138). The Taylor telehealth trained group improved from pretest (10.81) to posttest (12.41), as did the ACTIVE in-person trained group, increasing from pretest (11.29) to posttest (12.22). This can be seen in Figure 4-28.

Equivalency analysis

As a follow-up analysis, to determine if the Taylor telehealth trained and ACTIVE in-person trained groups had equivalent change in pretest to posttest performance on the 3 inductive reasoning tasks (Word Series, Letter Series, and Letter Sets), an equivalency test was conducted, See Figure 4-29. The confidence intervals around our obtained differences transects our 10% equivalence interval / zone of equivalent difference (see Table 4-7). The 95% Confidence Interval for Word Series (2.18 to 4.92) shares no overlap with the equivalency bounds of the 10% equivalence interval (-0.552 to 0.552), see Figure 4-30. The 95% Confidence Interval for Letter Series (-0.34 to 2.93) transects one of the equivalency bounds of the 10% equivalence interval (-0.446 to 0.446), see Figure 4-31. The 95% Confidence Interval for Letter Sets (-1.40 to 0.71) transects both equivalency bounds of the 10% equivalence interval (-0.098 to 0.098), see Figure 4-32. Therefore, we reject the null hypotheses of equivalence for all 3 inductive reasoning tasks. This likely reflects our lack of power, but we cannot claim the pretest to posttest performance changes were equivalent between the Taylor telehealth trained and the ACTIVE in-person trained.

Strategy use

All means and standard deviations for the number of correct and number of attempts on Word Series, Letter Series, and Letter Sets for Taylor Telehealth strategy users and non-strategy users are displayed in Table 4-9.

To determine whether annotated strategy use (underlining, slashes, and/or tick marks) affected the number of items answered correctly, telehealth-trained participants were divided into two groups (those who used any strategies at posttest and those who did not use any strategies at posttest) and were compared on their change in pretest to posttest inductive reasoning number of correct responses. As depicted in Figure 4-33, an ANOVA revealed a non-significant interaction between the effects of Occasion (pretest and posttest) and groups (strategy users and non-strategy users) on the number of correct items on Word Series [$F(1, 29) = 0.111$, $p = 0.741$, *partial eta-*

squared = 0.004]. As depicted in Figure 4-34, an ANOVA revealed a non-significant interaction between the effects of Occasion (pretest and posttest) and groups (strategy users and non-strategy users) on the number of correct items on Letter Series [$F(1, 29) = 1.896$, $p = 0.179$, *partial eta-squared* = 0.061]. As depicted in Figure 4-35, an ANOVA revealed a non-significant interaction between the effects of Occasion (pretest and posttest) and groups (strategy users and non-strategy users) on the number of correct items on Letter Sets [$F(1, 29) = 2.270$, $p = 0.143$, *partial eta-squared* = 0.073].

To determine whether annotated strategy use (underlining, slashes, and/or tick marks) affected the total number of items answered, telehealth-trained participants were divided into two groups (those who used any strategies at posttest and those who did not use any strategies at posttest) and were compared on their change in the pretest to posttest inductive reasoning total number of items answered. As depicted in Figure 4-36, an ANOVA revealed a non-significant interaction between the effects of Occasion (pretest and posttest) and groups (strategy users and non-strategy users) on the total number of attempted items on Word Series [$F(1, 29) = 0.644$, $p = 0.429$, *partial eta-squared* = 0.022]. As depicted in Figure 4-37, an ANOVA revealed a non-significant interaction between the effects of Occasion (pretest and posttest) and groups (strategy users and non-strategy users) on the total number of attempted items on Letter Series [$F(1, 29) = 0.790$, $p = 0.381$, *partial eta-squared* = 0.027]. As depicted in Figure 4-38, an ANOVA revealed a non-significant interaction between the effects of Occasion (pretest and posttest) and groups (strategy users and non-strategy users) on the total number of attempted items on Letter Sets [$F(1, 29) = 1.117$, $p = 0.299$, *partial eta-squared* = 0.037].

Overall, based on these results, we failed to find a difference between those who used strategies overtly vs. those who did not. However, some participants could have used internal strategies, which is something we encouraged late in training. Additionally, we lacked statistical power to detect a difference between these two small groups. Furthermore, the same ratio of people (20/31) used strategies on Word Series and Letter Series, while (15/31) used strategies on Letter Sets; therefore, this does not appear to explain the unique Word Series findings.

Training session-to-session curves

We wanted to explore whether the acquisition curves between the Taylor telehealth-trained participants and the ACTIVE in-person trained participants differed. This would help to inform us about the pattern or nature of improvement across the ten training sessions (i.e. if there is a little improvement after each session then that would represent incremental progress, whereby training builds upon the last session; however, as another example, if everyone obtains all their gains by training session 2 and then they plateau or no longer improve over training sessions 3-10, then that would suggest that perhaps only the first two training sessions are worthwhile and then participants are experiencing maintenance, ceiling effects of the inductive reasoning measures, or the training is no longer adding anything). To examine whether the Taylor telehealth-trained participants differed from the ACTIVE in-person trained participants on their session-to-session inductive reasoning performance acquisition curves, the end-of-session inductive reasoning number of correct items out of 20 and number of attempted items out of 20 were graphed (see Figures 4-39 and Figure 4-40).

Looking at the graphs (Figures 4-39 and Figure 4-40), it is quite clear that the overall story is that there was roughly parallel improvement between the training groups across the ten training sessions and both trainings were effective at creating performance improvement. However, we know from our prior work that in the ACTIVE in-person trained group, we found that training in that group occurred in 2 phases (initial fast gain, then slower ongoing gain), and that different variables significantly predict these slopes. Age, education, gender, and training exposure predicted initial fast gain, while education, training exposure, and MMSE scores predicted ongoing growth (Taylor et al., 2020). Because of this, we were surprised to see that the telehealth trained group had largely consistent linear change across the ten training sessions. However, the telehealth group took until session 4 to start improving, a stark contrast from the rapid growth in the initial sessions seen in the in-person trained group. It is possible that the telehealth trained participants miss out on an initial "a-ha" moment/faster initial growth because of the technological burden, but they were able to demonstrate steady inductive reasoning improvements as the sessions progressed from there. It is also worth noting that by the end of the study there was a pretty consistent 2-point gap between the groups that would have been statistically significant in a larger sample. We will discuss more in the discussion section, below.

Multi-Level Modeling was conducted to better examine the effects of group status (Taylor Telehealth vs ACTIVE in-person trained), session number (1-10), and their interaction on the number of correctly answered inductive reasoning problems across the sessions and the number of attempted inductive reasoning problems across the sessions on the 20-question end-of-session inductive reasoning assessments. These can be seen in Table 4-10 and Table 4-11. Intervention group status (Taylor Telehealth vs ACTIVE in-person trained) significantly predicted inductive reasoning number correct [$t(145.980) = 3.461$, $p < 0.001$], with individuals in the ACTIVE in-person trained group answering 2.76 more correct on average across the ten training sessions. There was a significant effect of session number (1-10) on inductive reasoning number correct [$t(146.527) = 9.581$, $p < 0.001$], such that participants improved by an average of 0.63 points as the sessions progressed. The interaction between groups (Taylor Telehealth vs ACTIVE in-person trained) by session (1-10) was non-significant [$t(146.527) = -1.705$, $p = 0.090$]. Intervention group status (Taylor Telehealth vs ACTIVE in-person trained) significantly predicted inductive reasoning number attempted [$t(138.193) = 3.887$, $p < 0.001$], with individuals in the ACTIVE in-person trained group attempting 2.72 more problems on average across the ten training sessions. There was a significant effect of session number (1-10) on inductive reasoning number attempted [$t(147.052) = 7.657$, $p < 0.001$], such that participants attempted an average of 0.52 additional questions as the sessions progressed. The interaction between groups (Taylor Telehealth vs ACTIVE in-person trained) by session (1-10) was not significant [$t(145.415) = -0.919$, $p = 0.360$].

Additional Qualitative Observations

Although the results section has focused primarily on quantitative evaluations of hypotheses and follow-up analyses, the experience of conducting the study also permitted observation of unexpected events that could not be captured by our quantitative measures. The brief section that follows includes additional qualitative observations about conducting Inductive Reasoning training via telehealth.

Extra Makeups/Flexible Scheduling

To accommodate the schedules of our participants, with the goal of maximizing adherence, flexible scheduling and make-up sessions were offered. When a participant informed study staff that they would be unable to attend one of their regularly scheduled visits, the first step was to first see if they were able to join another training group's equivalent session number at a different time. If that did not work their schedule, then individual make-up sessions were scheduled and occasionally combined with other's make-up sessions. Sessions were also scheduled on a very flexible basis, with regular sessions occurring on evenings and weekends, in addition to normal Monday-Friday business hours, to accommodate the consensus of training groups as they were formed.

This flexibility differed substantially from the ACTIVE trial, which was much more regimented with their schedules. ACTIVE did not offer evening or weekend training sessions (although they did offer regular make-up sessions on an as-needed basis). These study differences in flexibility might have an unknown effect on group comparability and retention, although retention was high in both studies. It is important to note that telehealth afforded more flexibility (and this may in fact constitute one of the main benefits of telehealth training paradigms) because no scheduling constraints were imposed by fixed schedules for rented/booked rooms or "standard office hours". Removing transportation and parking barriers likely also made participants more willing to consider non-business hours and days.

Technology

Technology, while being the primary point of difference from the ACTIVE trial, also had effects on some testing and training procedures. Firstly, during the MMSE's orientation questions, a few participants mentioned that they were able to find the date on the computer screen. As the study progressed, study staff told participants to not look at their computer for assistance with answering questions pertaining to the date; however, this still relied on the honor system, and differed from the ACTIVE traditional paper and pencil administration.

By relying on participant-provided equipment, the present study could not standardize personal computers, mice, and internet speeds; these were not sources of individual differences when ACTIVE employed a paper-pencil format for training workbooks. A better resourced study could have provided equipment to standardize these details, but the current study is probably more useful as a real-world test of telehealth training under natural conditional.

Slow internet speeds or intermittent wi-fi connections negatively impacted some participants. The intervention was also limited by Canvas problems, including one training session when Canvas was down, and the training staff had to pivot to screensharing Word documents and doing all the activities as a group. Canvas' annotated assignments' function had numerous glitches that occasionally made opening annotated assignments cumbersome, requiring the use of the refresh button. The unreliability of the Canvas system was interesting to note (and corroborated by instructors who were teaching at similar times), and it is somewhat surprising that such a widely deployed learning management system would be so error prone.

As noted in the results above, unlike two other inductive reasoning measures, Word Series findings under telehealth differed from what had been reported in ACTIVE. In the present study, telehealth-trained participants evinced significantly smaller improvements in Word Series than propensity-matched ACTIVE inductive reasoning trained participants; indeed, the Word Series improvement in telehealth participants was not significantly greater than that shown by ACTIVE no-contact controls. We suspect again that Zoom and the Canvas portal may have been responsible for this. From a display perspective, in our telehealth interface, each Word Series task took up an entire screen, increasing the number of page scrolls participants needed to make (relative to ACTIVE); we also suspect the mechanical time of Canvas page scrolls might have been longer than flipping a test booklet page (as was done in ACTIVE). This was not the case for Letter Series and Letter Sets, which in both the online and paper-and-pencil formats had multiple questions per page, meaning less time was spent scrolling/page flipping.

Use of Proctors/Aides in Session

One major, but necessary, deviation from the standard ACTIVE protocol was the use of proctors/research assistants in the training sessions to assist with the telehealth delivery. In the original ACTIVE trial, each training session was typically led by a single trainer who managed all 4-5 participants. However, in the present study, given the anticipated technological concerns that might be encountered by our older adult participants, we opted to assign at least one proctor/aid/research assistant. These assistants were undergraduate research assistants who had demonstrated proficiency in Zoom and Canvas. The assistants were included to assist participants with technology problems, freeing up the trainer to focus on lesson progress. In practice, we found that these research assistants needed to assist frequently, particularly in early sessions. This is a possible source of difference between ACTIVE and the present study. In addition, given how heavily used these assistants were, it suggests that future telehealth training might want to consider smaller group sizes or built-in supporters, at least during the early sessions.

Table 4-1. Means and standard deviations for Word Series, Letter Series, Letter Sets, Vocabulary, Self-Efficacy, and Concern about Aging for Taylor Telehealth, ACTIVE controls, and ACTIVE in-person.

		Taylor Telehealth			ACTIVE Controls			ACTIVE In-Person		
		N	mean	SD	*N*	mean	SD	*N*	mean	SD
Word Series	Pretest	31	10.9	3.44	87	11.49	5.24	87	10.47	4.94
	Posttest	31	12.87	3.82	87	14.11	5.1	87	15.99	5.60
Letter Series	Pretest	31	14.06	5.38	87	14.4	5.87	87	13.49	6.29
	Posttest	31	17.65	4.76	87	15.78	5.91	87	17.95	5.90
Letter Sets	Pretest	31	7.77	2.58	87	8.25	2.68	87	7.83	2.72
	Posttest	31	9.35	2.78	87	8.47	2.98	87	8.8	3.01
Vocabulary	Pretest	31	12.81	3.49	75	13.57	3.48	78	12.76	4.00
	Posttest / 1-Year Followup	31	13.52	2.89	75	13.95	2.96	78	13.69	3.23
Self Efficacy	Pretest	31	32.11	3.22	75	32.48	2.86	78	32.51	3.11
	Posttest / 1-Year Followup	31	33.58	2.66	75	32.24	3.49	78	32.56	3.03
Concern about Aging	Pretest	31	3.23	4.62	75	2.56	4.88	78	3.57	5.66
	Posttest/1-Year Followup	31	2.27	3.87	75	3.38	5.77	78	3.23	5.99

Note: N=group size, SD=standard deviation

Table 4-2. ANOVA examining the effects of Group (Taylor telehealth trained and ACTIVE controls), Time/Occasion (pretest and posttest), and the Group by Time interaction on Word Series, Letter Series, and Letter Sets scores.

	Word Series					Letter Series					Letter Sets				
Effect	df_n	df_d	*F*	*p*	η^2	df_n	df_d	*F*	*p*	η^2	df_n	df_d	*F*	*p*	η^2
G	1	116	0.91	0.342	0.01	1	116	0.45	0.504	0.004	1	116	0.14	0.707	0.001
T	1	116	56.53	< 0.001	0.33	1	116	49.79	< 0.001	0.30	1	116	15.15	< 0.001	0.12
G*T	1	116	1.15	0.287	0.01	1	116	9.81	0.002	0.08	1	116	8.69	0.004	0.07

Note: G = Group, T = Time/Occasion, G*T = Group by Time interaction, *dfn* = numerator degrees of freedom, *dfd* = denominator degrees of freedom, *F* = F statistic, *p* = probability of F statistic under null hypothesis, η^2 = partial eta squared (effect size).

Table 4-3. ANOVA examining the effects of Group (Taylor telehealth trained and ACTIVE controls), Time (pretest and posttest), and the Group by Time interaction on Vocabulary, Self-Efficacy, and Concern about Aging.

	Vocabulary					Self Efficacy					Concern About Aging				
Effect	df_n	df_d	F	p	η^2	df_n	df_d	F	p	η^2	df_n	df_d	F	p	η^2
Group	1	104	0.830	0.364	0.008	1	104	0.624	0.432	0.006	1	104	0.045	0.833	< 0.001
Time	1	104	7.042	0.009	0.063	1	104	5.773	0.018	0.053	1	104	0.036	0.851	< 0.001
G*T	1	104	0.679	0.412	0.006	1	104	11.169	0.001	0.097	1	104	6.303	0.014	0.057

Note: G*T = Group by Time interaction, *dfn* = numerator degrees of freedom, *dfd* = denominator degrees of freedom, F = F statistic, p = probability of F statistic under null hypothesis, η^2 = partial eta squared (effect size).

Table 4-4. Means and standard deviations for the number of attempts on Word Series, Letter Series, and Letter Sets for Taylor Telehealth, ACTIVE Controls, and ACTIVE In-Person.

		Taylor Telehealth			ACTIVE Controls			ACTIVE In-Person		
		N	mean	SD	*N*	mean	SD	*N*	mean	SD
Word Series	Pretest	31	13.29	3.50	87	14.35	5.02	87	14.41	4.03
	Posttest	31	13.90	3.90	87	16.94	4.83	87	18.74	5.58
Letter Series	Pretest	31	16.97	4.73	87	18.36	5.13	87	18.64	5.13
	Posttest	31	18.97	4.41	87	20.02	5.22	87	20.86	5.56
Letter Sets	Pretest	31	10.81	2.56	87	11.30	2.69	87	11.29	2.57
	Posttest	31	12.42	2.03	87	11.78	2.48	87	12.22	2.59

Note: N = group size, SD = standard deviation.

Table 4-5. ANOVA examining the effects of Group (Taylor telehealth trained and ACTIVE controls), Time (pretest and posttest), and the Group by Time interaction on the number of attempted problems on Word Series, Letter Series, and Letter Sets.

	Number of Word Series Attempts					Number of Letter Series Attempts					Number of Letter Sets Attempts				
Effect	df_n	df_d	F	p	η^2	df_n	df_d	F	p	η^2	df_n	df_d	F	p	η^2
Group	1	116	5.265	0.024	0.4	1	116	1.61	0.207	0.01	1	116	0.25	0.874	0.00
Time	1	116	17.38	<0.005	0.13	1	116	27.80	<0.001	0.19	1	116	15.88	<0.001	0.12
Group * Time	1	116	6.64	0.011	0.05	1	116	0.14	0.710	0.00	1	116	4.62	0.034	0.04

Note: *dfn* = numerator degrees of freedom, *dfd* = denominator degrees of freedom, *F* = F statistic, *p* = probability of F statistic under null hypothesis, η^2 = partial eta squared (effect size).

Table 4-6. ANOVA examining the effects of Group (Taylor telehealth trained and ACTIVE Reasoning), Time (pretest and posttest), and the Group by Time interaction on Word Series, Letter Series and Letter Sets scores.

	Word Series					Letter Series					Letter Sets				
Effect	df_n	df_d	F	p	η^2	df_n	df_d	F	p	η^2	df_n	df_d	F	p	η^2
Group	1	116	1.94	0.167	0.02	1	116	0.01	0.910	< 0.001	1	116	0.22	0.638	0.002
Time	1	116	114.20	< 0.001	0.50	1	116	104.44	< 0.001	0.47	1	116	23.14	< 0.001	0.17
Group * Time	1	116	25.68	< 0.001	0.18	1	116	1.25	0.226	0.01	1	116	1.29	0.259	0.01

Note: *dfn* = numerator degrees of freedom, *dfd* = denominator degrees of freedom, *F* = F statistic, *p* = probability of F statistic under null hypothesis, η^2 = partial eta squared (effect size).

Table 4-7. ANOVA examining the effects of Group (Taylor telehealth trained and ACTIVE Reasoning), Time (pretest and posttest), and the Group by Time interaction on Vocabulary, Self-Efficacy, and Concern about Aging.

	Vocabulary					Self Efficacy					Concern About Aging				
Effect	df_n	df_d	*F*	*p*	η^2	df_n	df_d	*F*	*p*	η^2	df_n	df_d	*F*	*p*	η^2
Group	1	107	0.01	0.929	0.00	1	107	0.29	0.593	0.00	1	107	0.35	0.558	0.00
Time	1	107	11.75	< 0.001	0.10	1	107	6.78	0.011	0.06	1	107	3.45	0.066	0.03
Group * Time	1	107	0.22	0.638	0.00	1	107	5.90	0.017	0.05	1	107	0.77	0.384	0.01

Note: *dfn* = numerator degrees of freedom, *dfd* = denominator degrees of freedom, *F* = F statistic, *p* = probability of F statistic under null hypothesis, η^2 = partial eta squared (effect size).

Table 4-8. ANOVA examining the effects of Group (Taylor telehealth trained and ACTIVE Reasoning), Time (pretest and posttest), and the Group by Time interaction on the number of attempted problems on Word Series, Letter Series, and Letter Sets.

	Number of Word Series Attempts					Number of Letter Series Attempts					Number of Letter Sets Attempts				
Effect	df_n	df_d	*F*	*p*	η^2	df_n	df_d	*F*	*p*	η^2	df_n	df_d	*F*	*p*	η^2
Group	1	116	12.11	<0.001	0.10	1	116	3.17	0.077	0.03	1	116	0.10	0.748	0.00
Time	1	116	32.02	< 0.001	0.22	1	116	27.80	<0.001	0.19	1	116	18.51	<0.001	0.14
Group * Time	1	116	18.09	< 0.001	0.14	1	116	0.75	0.785	0.00	1	116	1.33	0.251	0.01

Note: *dfn* = numerator degrees of freedom, *dfd* = denominator degrees of freedom, *F* = F statistic, *p* = probability of F statistic under null hypothesis, η^2 = partial eta squared (effect size).

Table 4-9. Means and standard deviations for the number of correct and number of attempts on Word Series, Letter Series, and Letter Sets for Taylor Telehealth strategy users and non-strategy users.

		Taylor Telehealth Strategy Users			Taylor Telehealth Non-Strategy Users		
		N	mean	SD	*N*	mean	SD
Word Series # Correct	Pretest	20	10.55	3.33	11	11.55	3.70
	Posttest	20	12.40	4.17	11	13.73	3.07
Letter Series # Correct	Pretest	20	13.85	5.72	11	14.45	4.93
	Posttest	20	16.75	5.12	11	19.27	3.69
Letter Sets # Correct	Pretest	15	7.00	2.75	16	8.50	2.25
	Posttest	15	9.20	3.19	16	9.50	2.42
Word Series # Attempted	Pretest	20	13.00	3.61	11	13.82	3.37
	Posttest	20	13.35	4.08	11	14.91	3.51
Letter Series # Attempted	Pretest	20	16.45	4.81	11	17.91	4.66
	Posttest	20	18.05	4.75	11	20.64	3.26
Letter Sets # Attempted	Pretest	15	10.00	3.14	16	11.56	1.63
	Posttest	15	12.07	2.55	16	12.75	1.39

Note: N = group size, SD = standard deviation.

Table 4-10. MLM effects of group status (Taylor Telehealth vs ACTIVE in-person trained), session number (1-10), and their interaction on the number of correctly answered inductive reasoning problems across the sessions.

Parameter	Estimate	Std. Error	*df*	*t*	*p*	95% CI Lower Bound	95% CI Upper Bound
Intercept	7.25	0.69	146.93	10.56	< 0.001	5.89	8.61
ACTIVE Reasoning	2.76	0.80	145.98	3.46	< 0.001	1.18	4.34
Session #	0.63	0.07	147.78	9.58	< 0.001	0.50	0.76
Group * Session #	-0.13	0.08	146.53	-1.70	0.090	-0.28	0.02

Note: Estimate = unstandardized regression weight, Std. Error = Standard Error, *df* = degrees of freedom, *t* = t statistic (estimate/standard error), *p* = probability of t-statistic under null hypothesis, CI = 95% confidence interval.

Table 4-11. MLM effects of group status (Taylor Telehealth vs ACTIVE in-person trained), session number (1-10), and their interaction on the number of attempted inductive reasoning problems across the sessions.

Parameter	Estimate	Std. Error	*df*	*t*	*p*	95% CI Lower Bound	95% CI Upper Bound
Intercept	8.52	0.60	140.01	14.11	< 0. 001	7.33	9.72
ACTIVE Reasoning	2.72	0.70	138.19	3.89	< 0. 001	1.34	4.11
Session #	0.52	0.07	147.05	7.66	< 0. 001	0.38	0.65
Group * Session #	-0.07	0.08	145.42	-0.92	0.360	-0.23	0.08

Note: Estimate = unstandardized regression weight, Std. Error = Standard Error, *df* = degrees of freedom, *t* = t statistic (estimate/standard error), *p* = probability of t-statistic under null hypothesis, CI = 95% confidence interval.

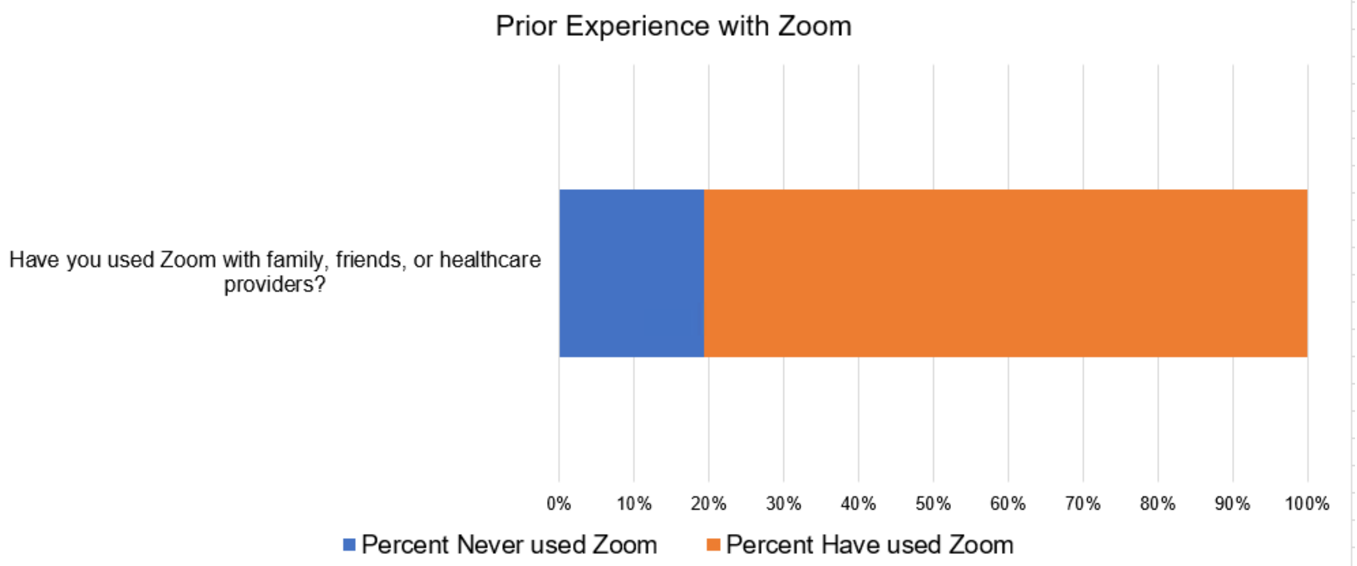


Figure 4-1. Prior Experience with Zoom.

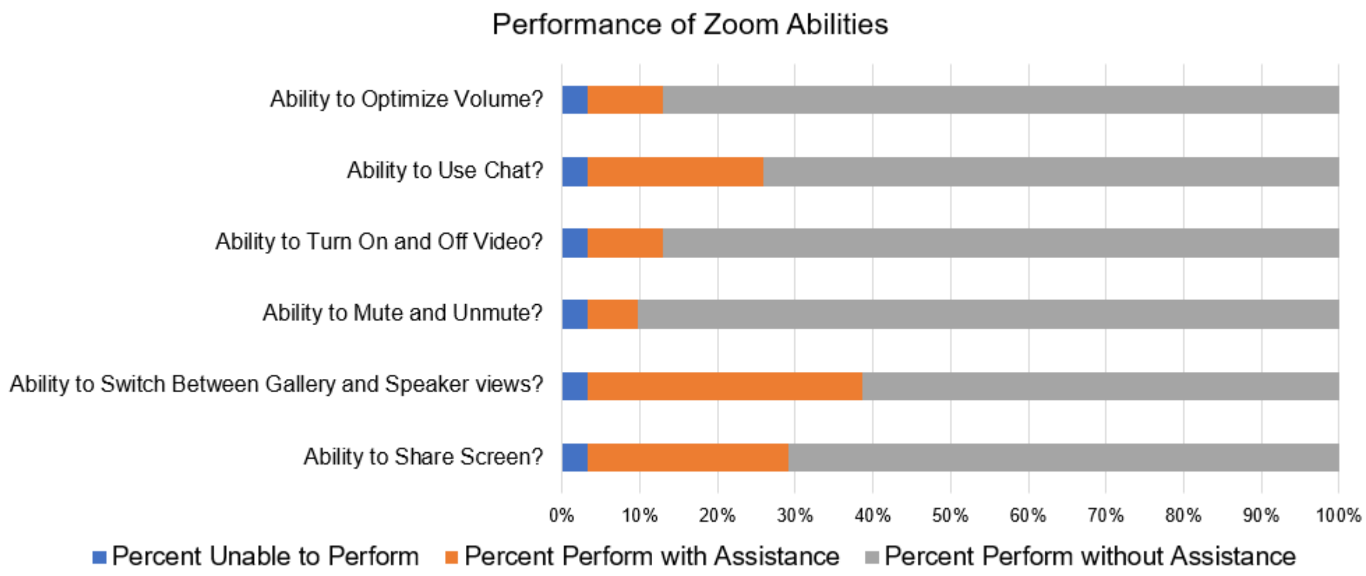


Figure 4-2. Performance of Zoom Abilities.

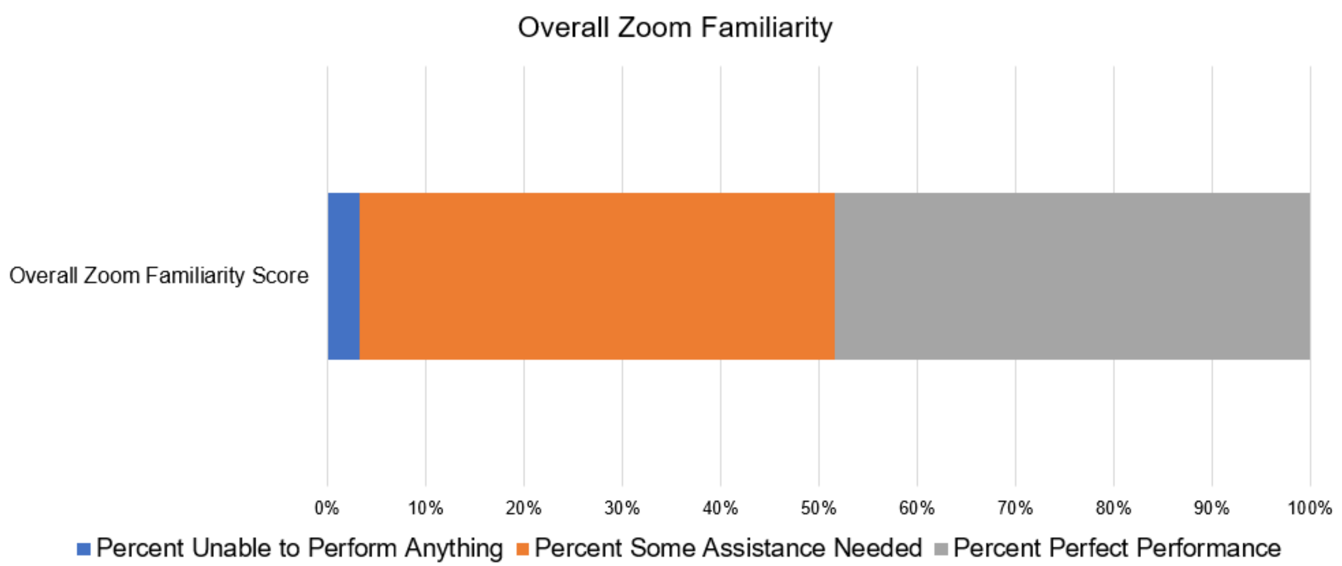


Figure 4-3. Overall Zoom Familiarity.

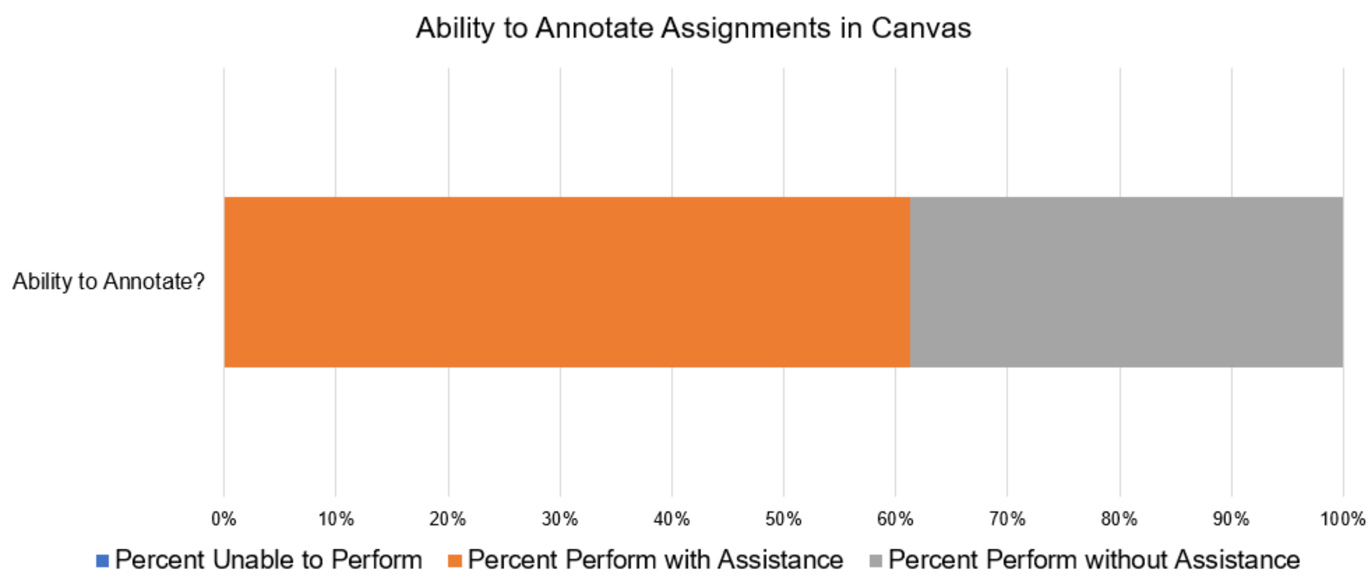


Figure 4-4. Ability to Annotate Assignments in Canvas.

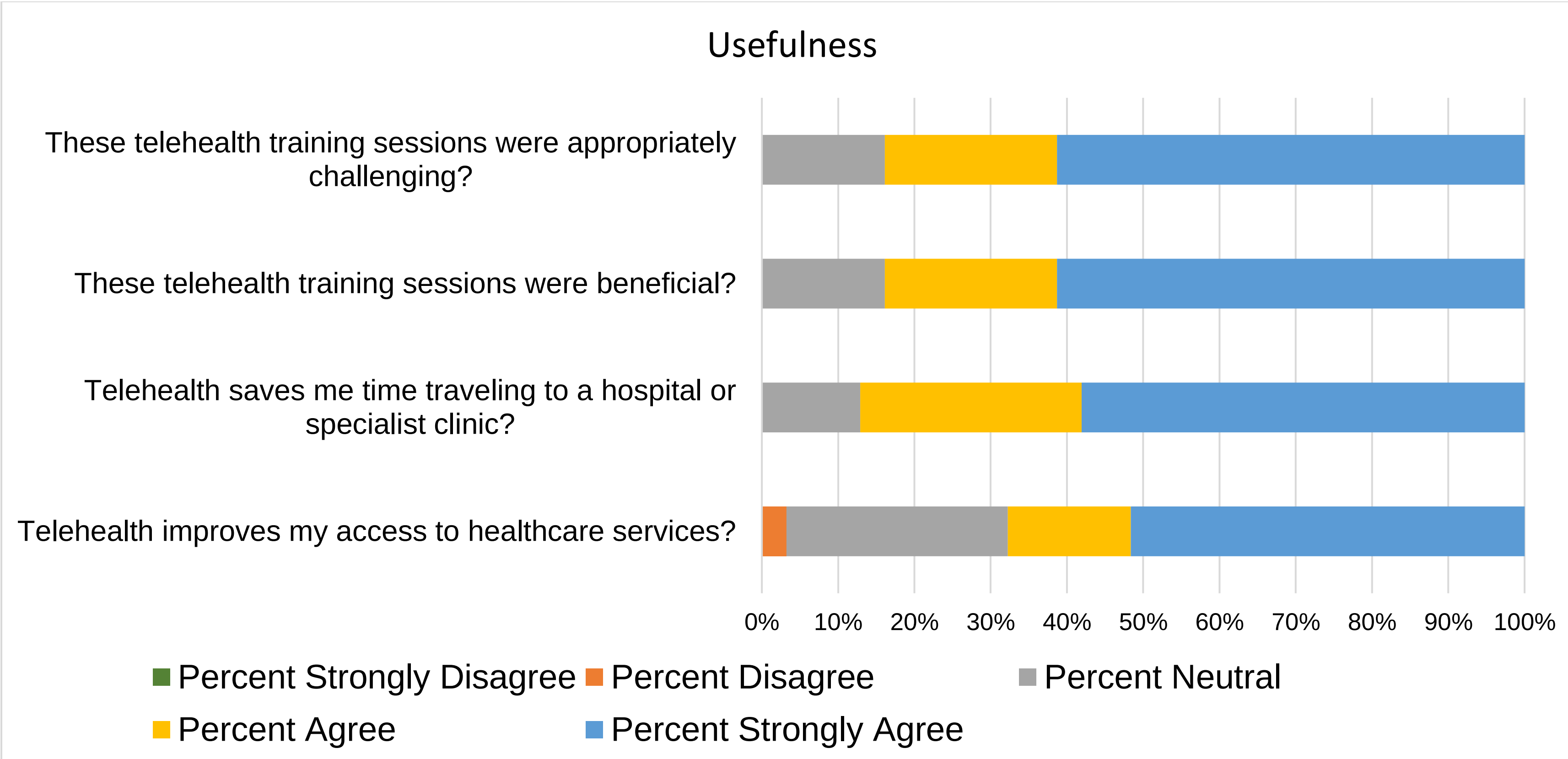


Figure 4-5. Usefulness.

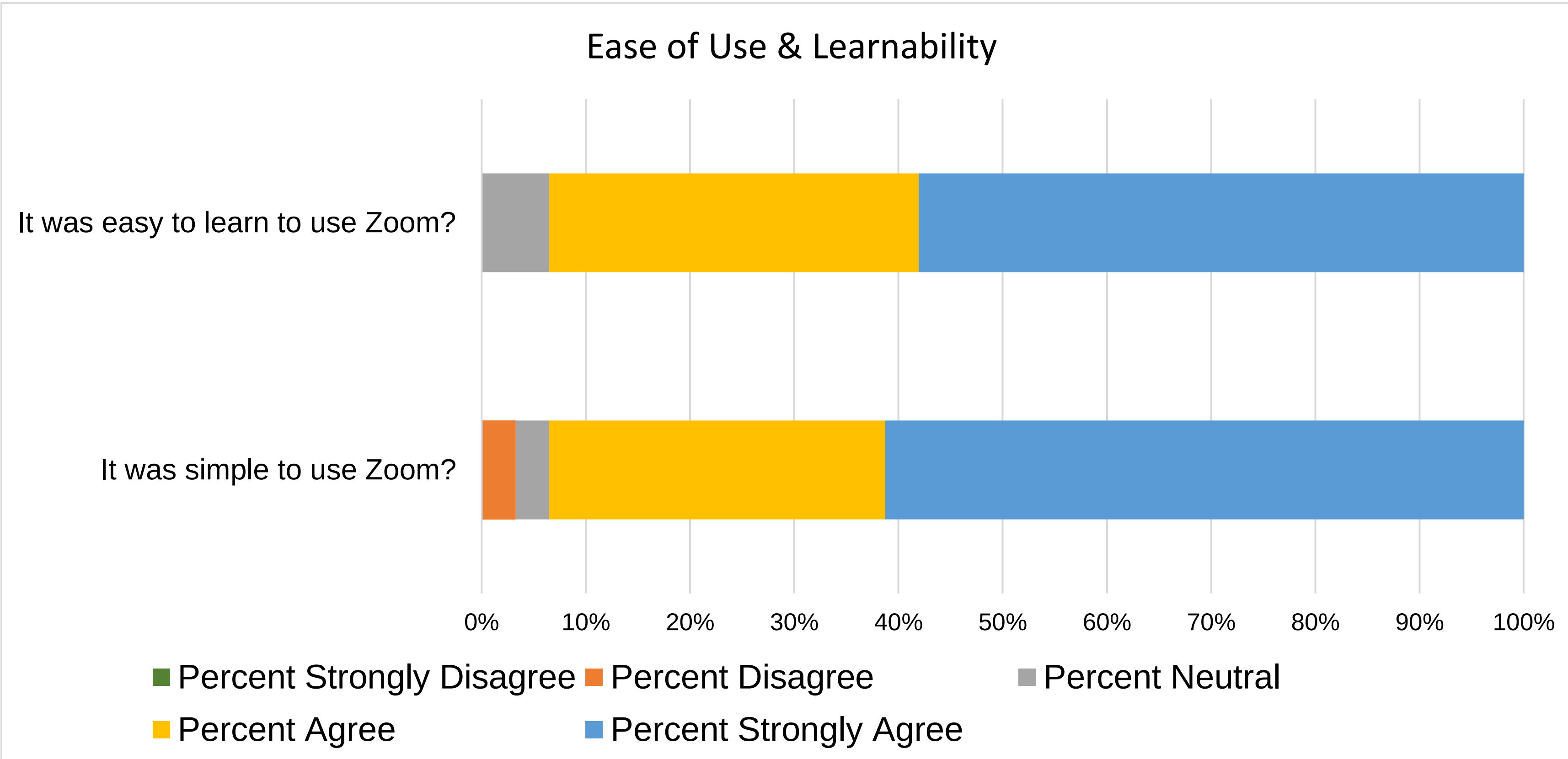


Figure 4-6. Ease of Use & Learnability.

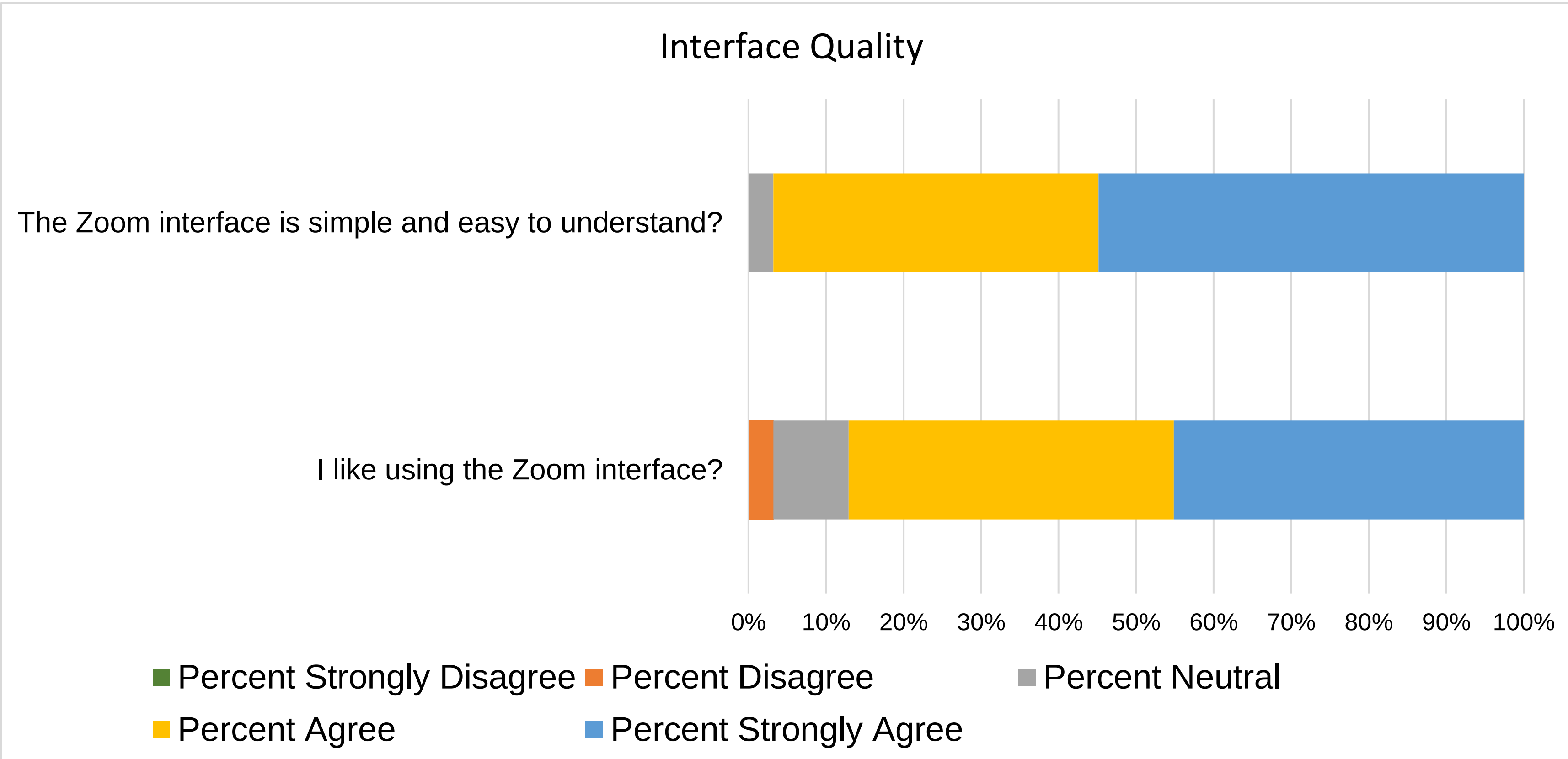


Figure 4-7. Interface Quality.

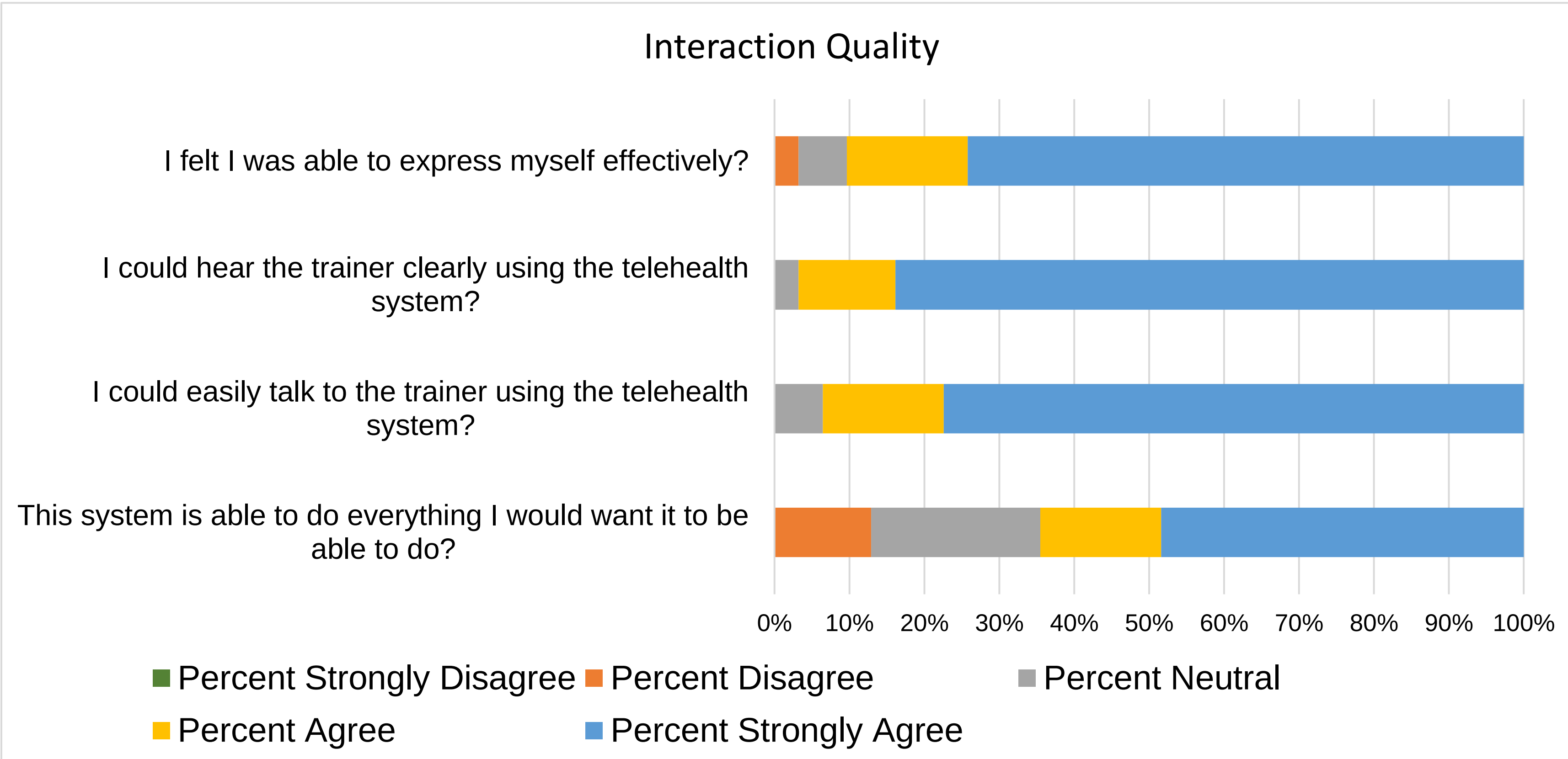


Figure 4-8. Interaction Quality.

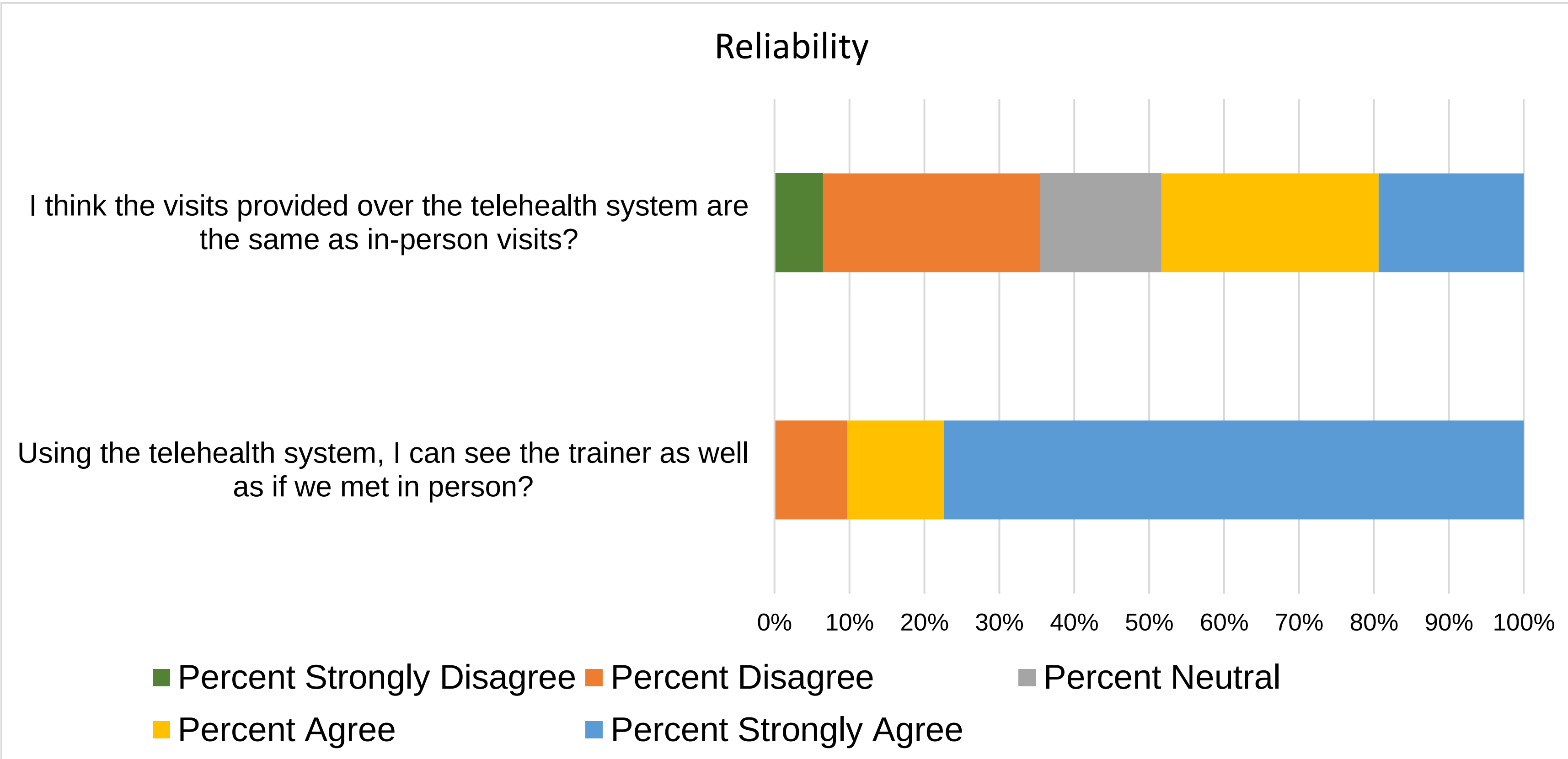


Figure 4-9. Reliability.

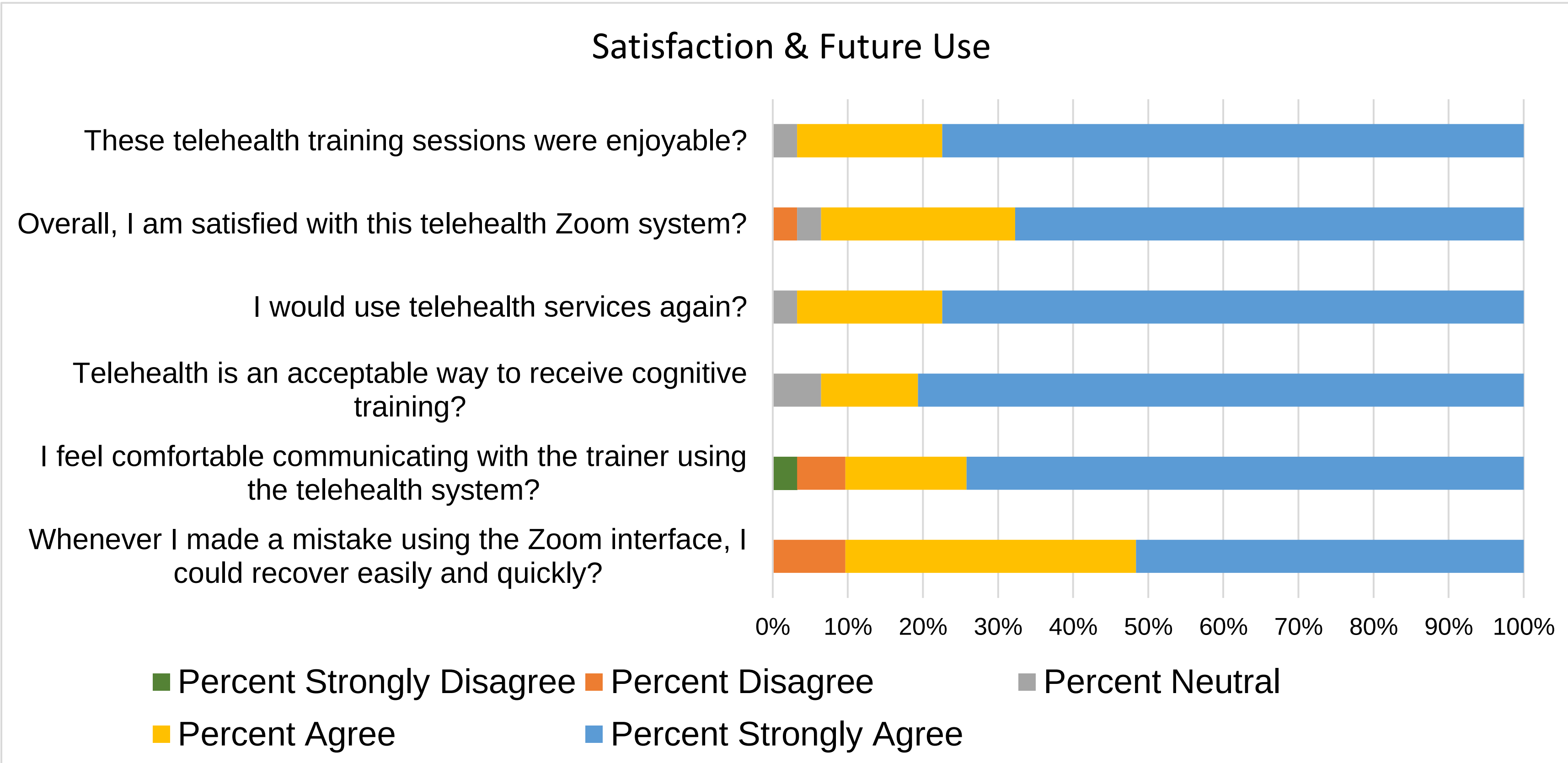


Figure 4-10. Satisfaction & Future Use.

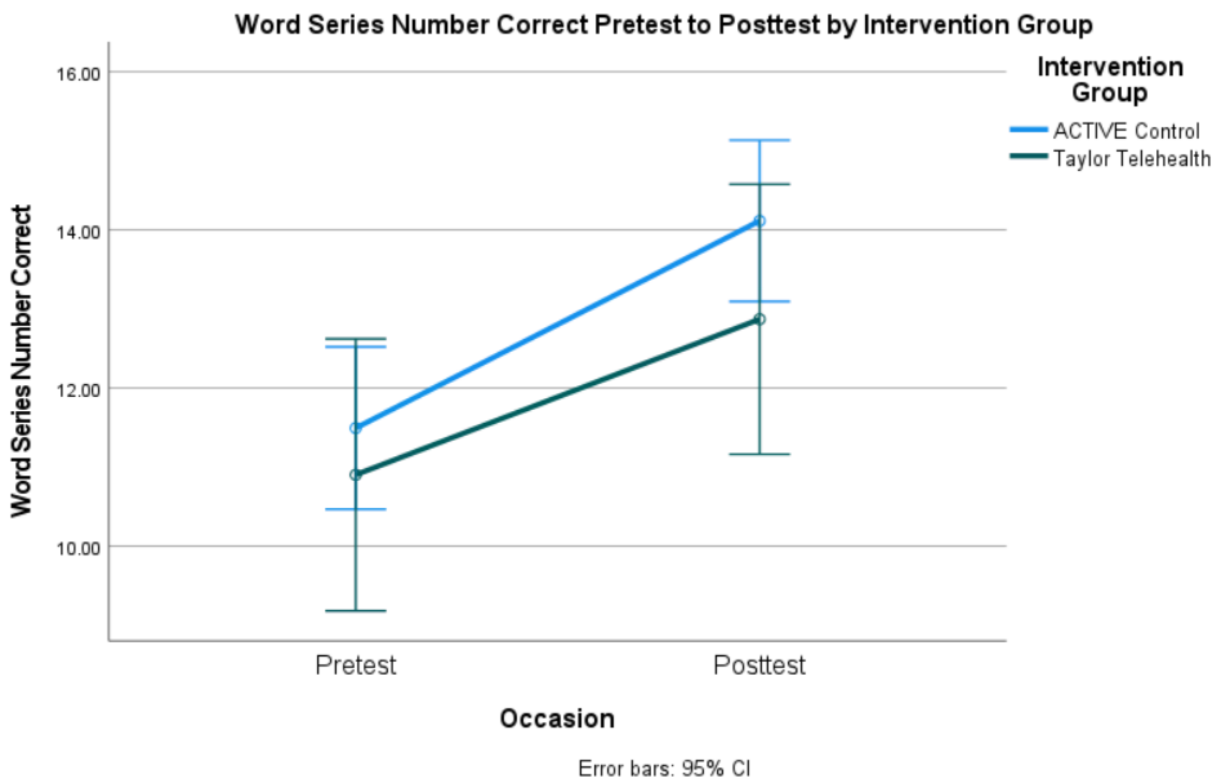


Figure 4-11. Word Series Number Correct Pretest to Posttest by Intervention Group.

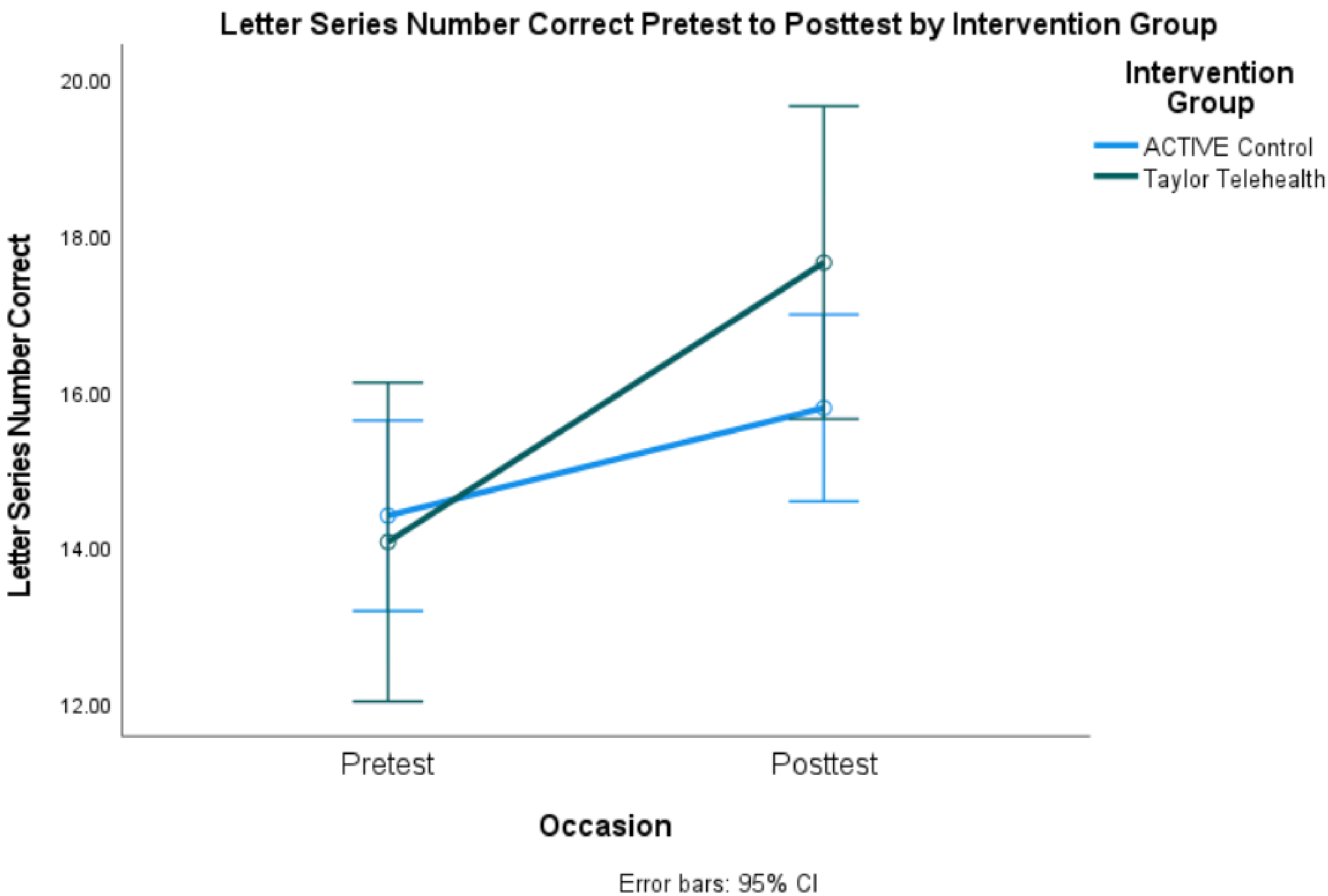


Figure 4-12. Letter Series Number Correct Pretest to Posttest by Intervention Group.

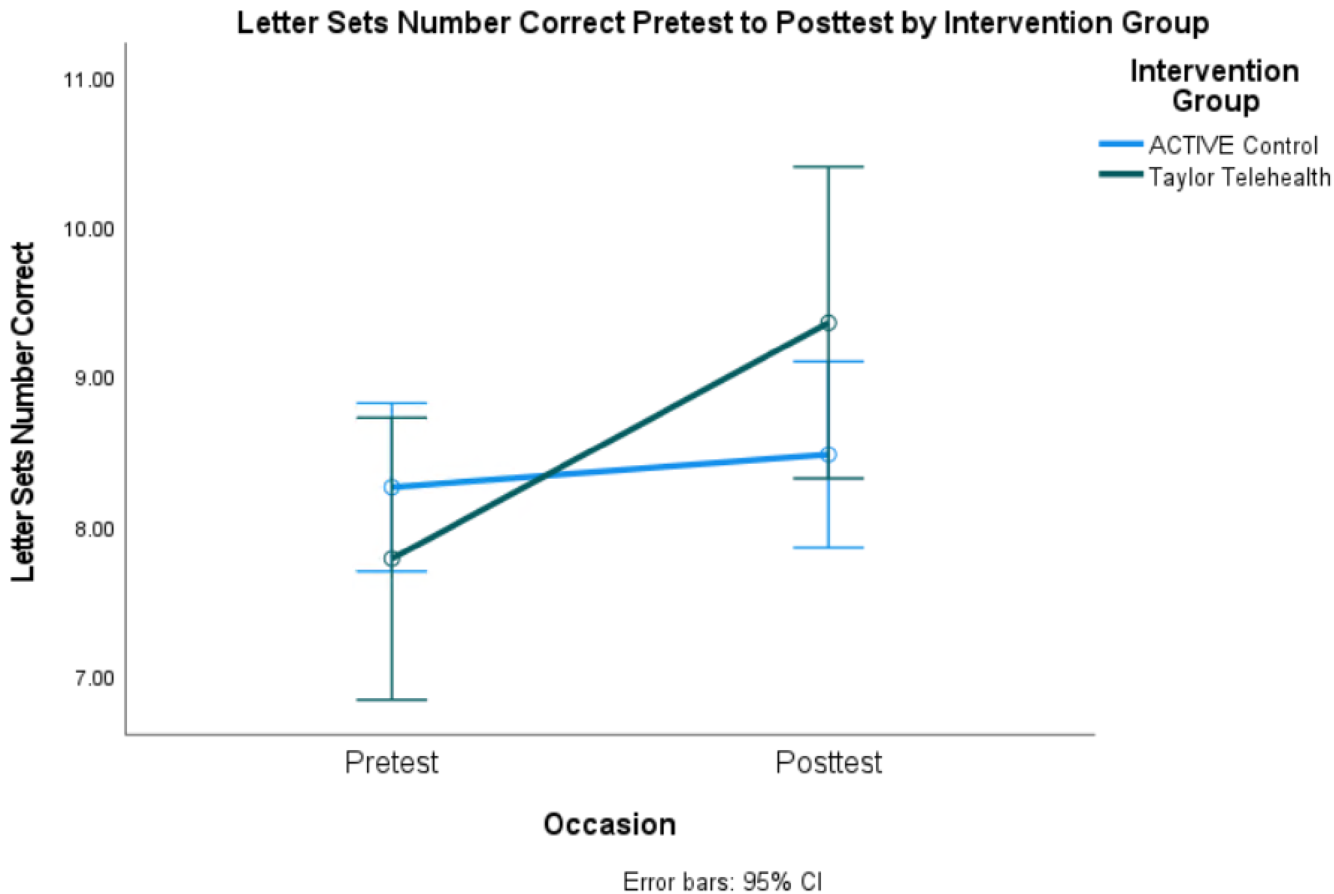


Figure 4-13. Letter Sets Number Correct Pretest to Posttest by Intervention Group.

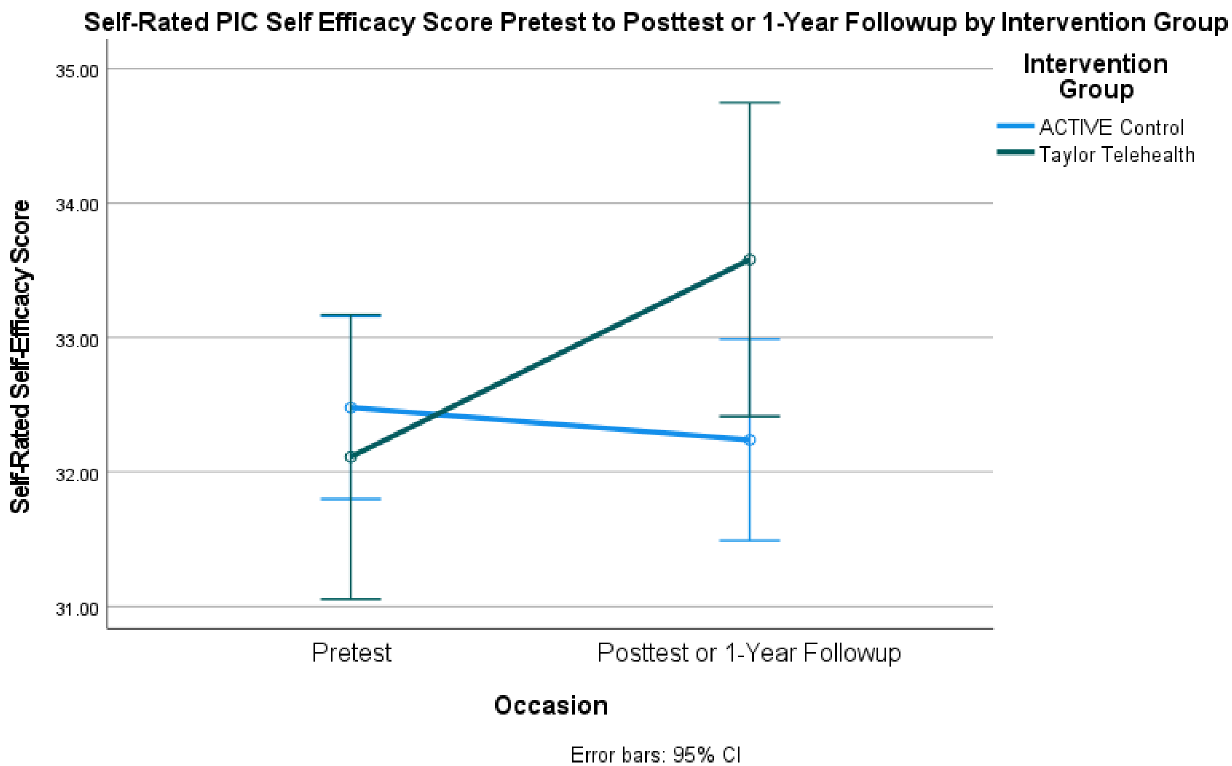


Figure 4-14. Self-Rated PIC Self Efficacy Score Pretest to Posttest or 1-Year Followup by Intervention Group.

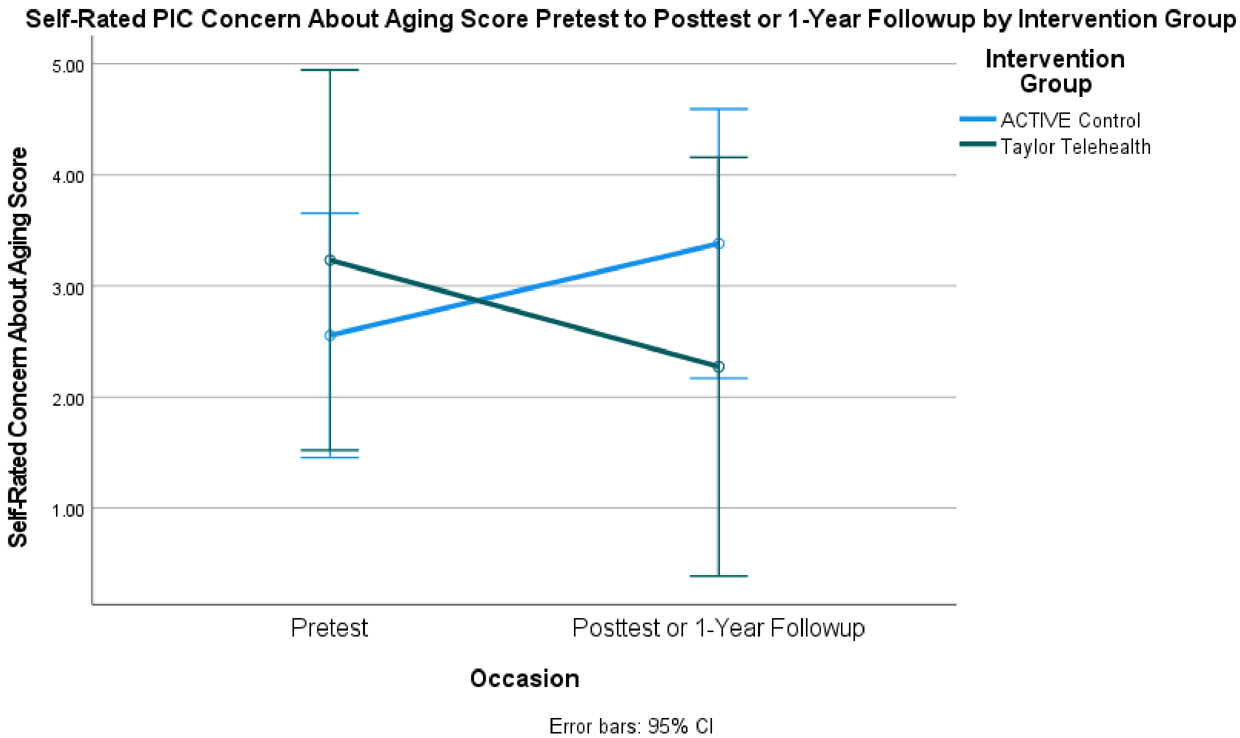


Figure 4-15. Self-Rated PIC Concern About Aging Score Pretest to Posttest or 1-Year Followup by Intervention Group.

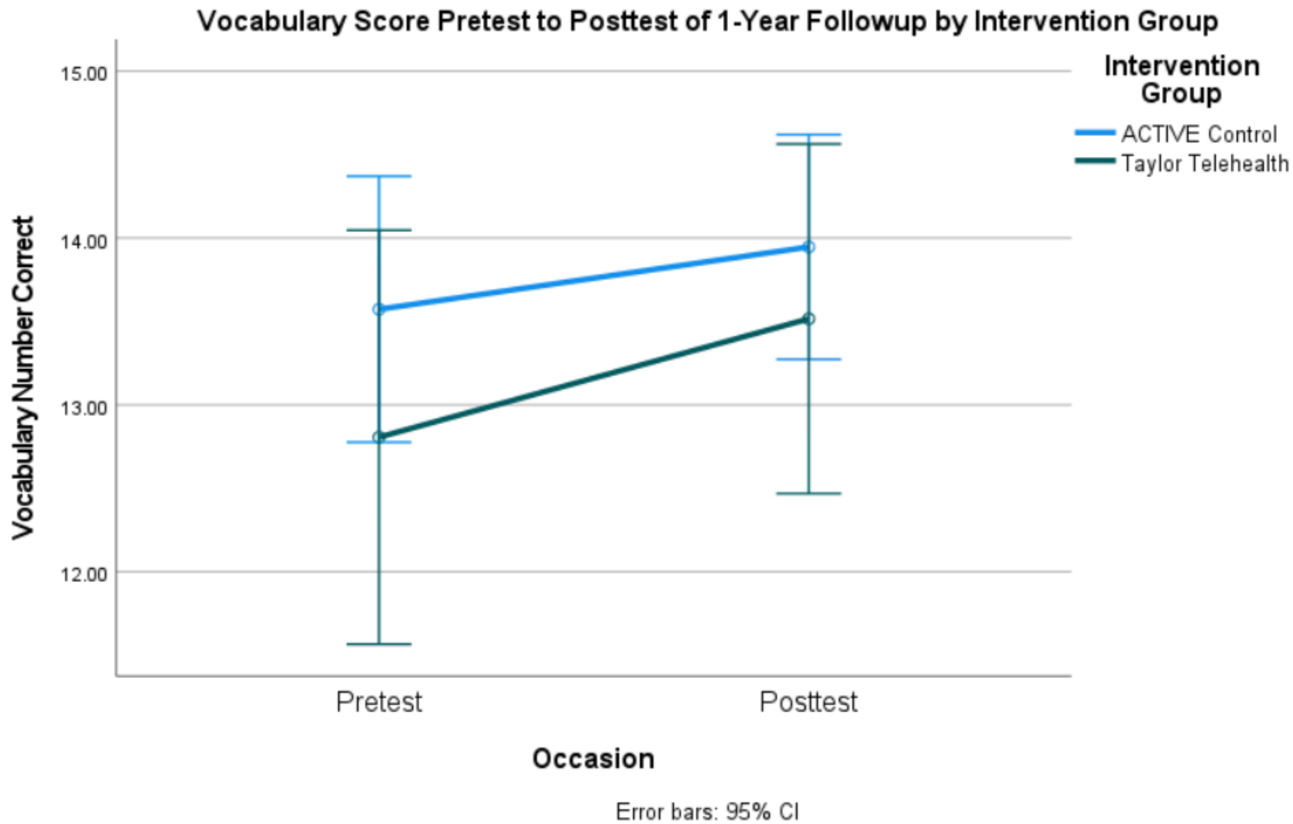


Figure 4-16. Vocabulary Score Pretest to Posttest or 1-Year Followup by Intervention Group.

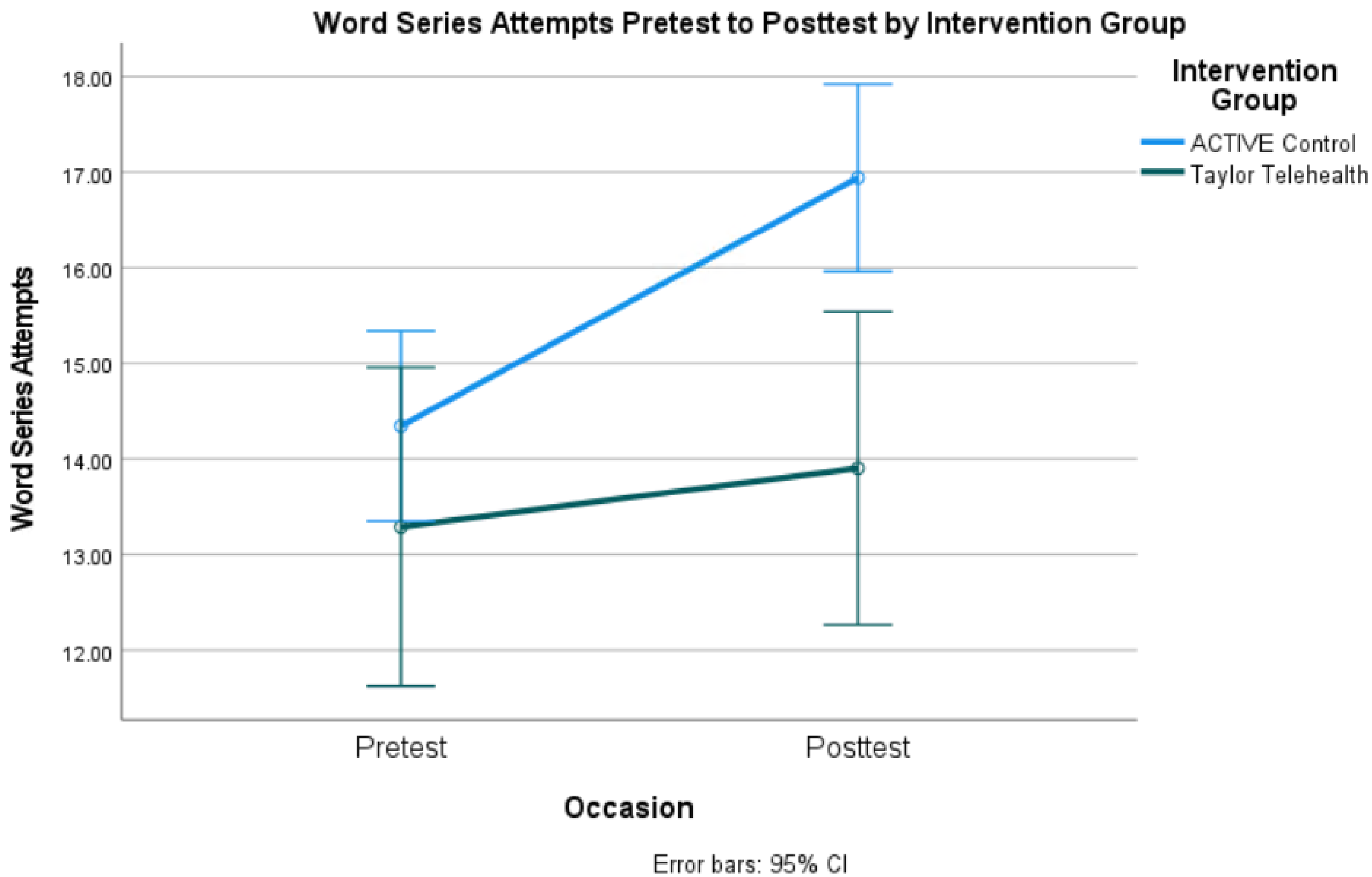


Figure 4-17. Word Series Attempts Pretest to Posttest by Intervention Group.

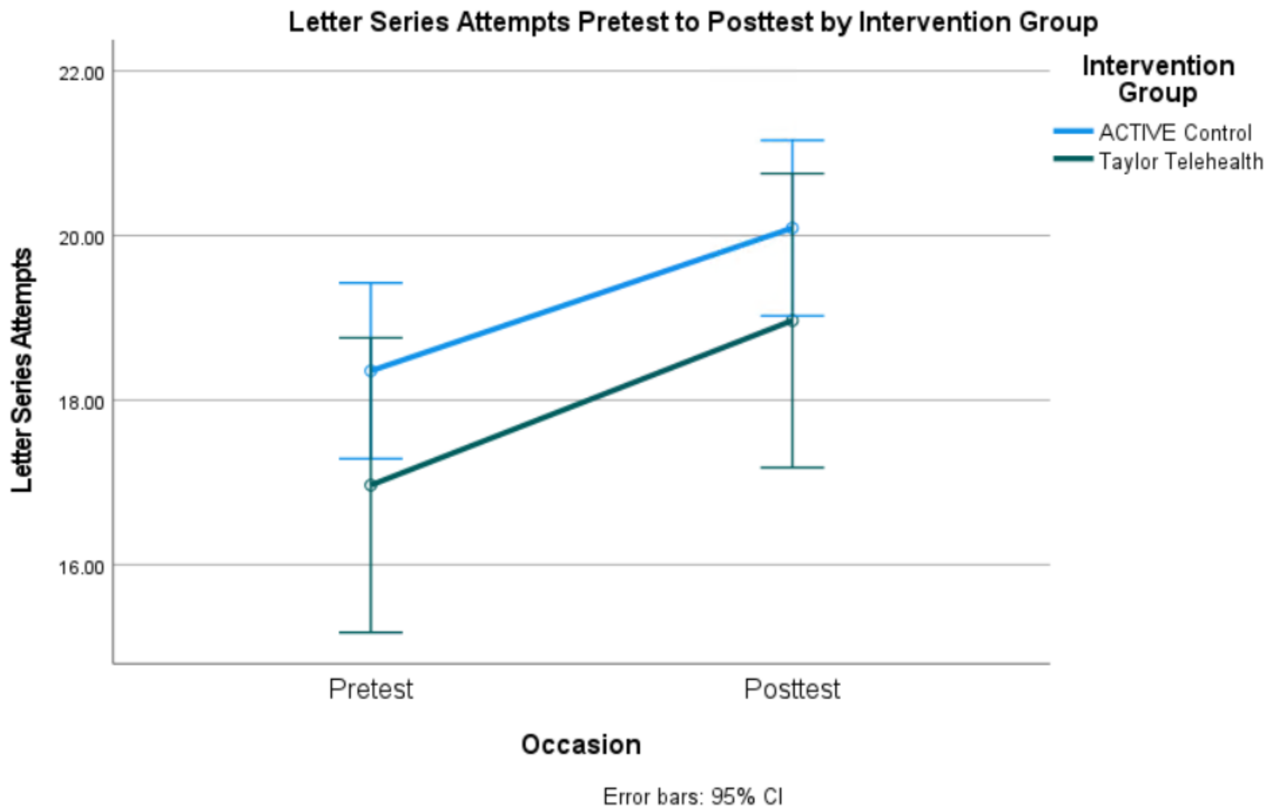


Figure 4-18. Letter Series Attempts Pretest to Posttest by Intervention Group.

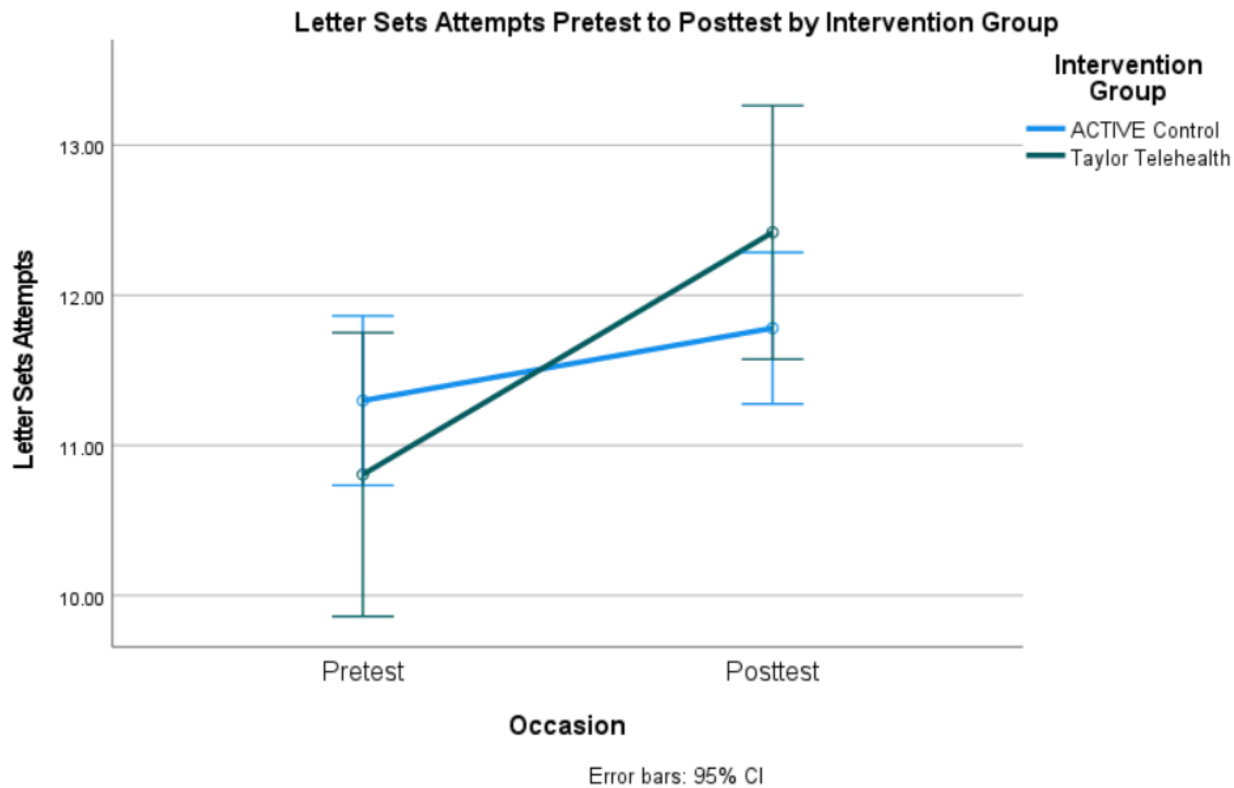


Figure 4-19. Letter Sets Attempts Pretest to Posttest by Intervention Group.

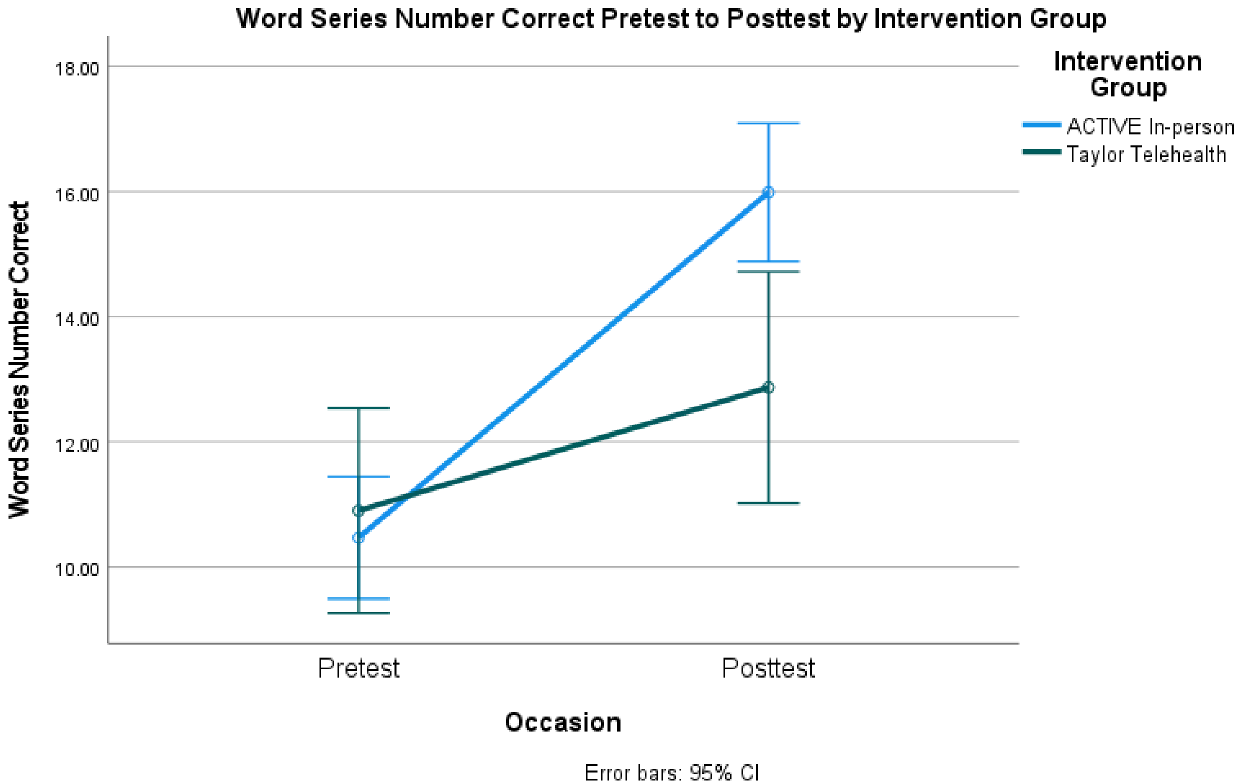


Figure 4-20. Word Series Number Correct Pretest to Posttest by Intervention Group.

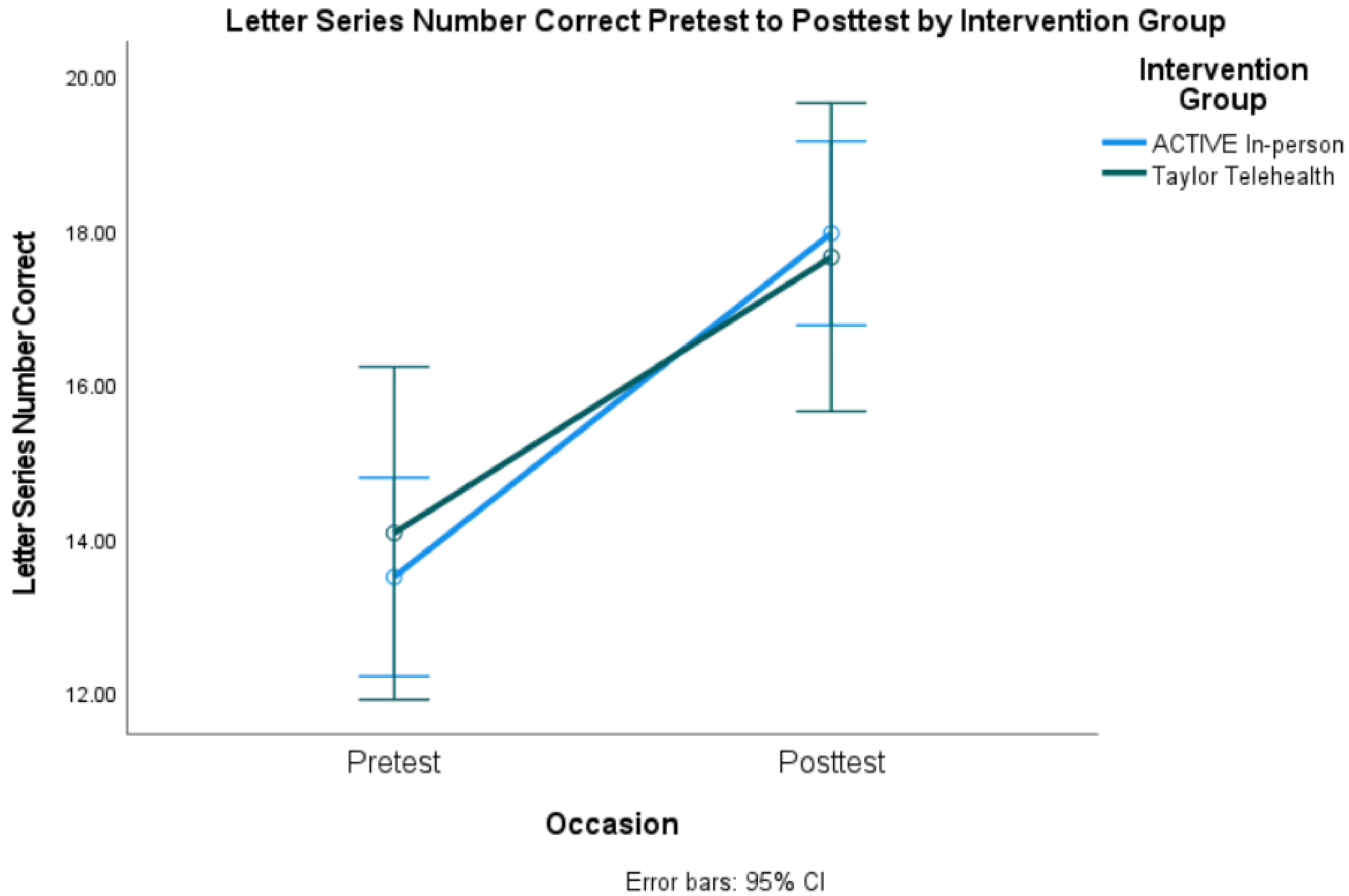


Figure 4-21. Letter Series Number Correct Pretest to Posttest by Intervention Group.

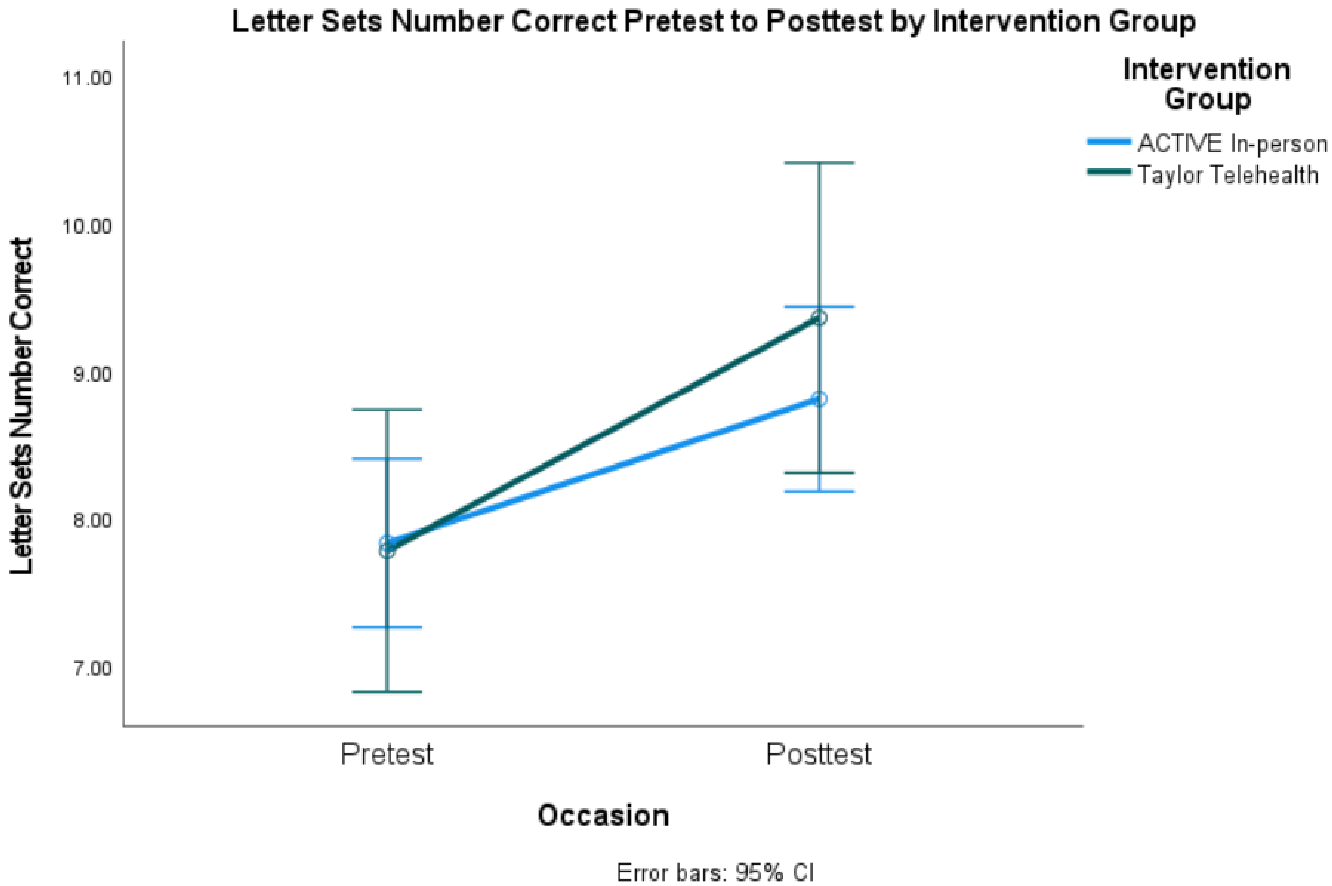


Figure 4-22. Letter Sets Number Correct Pretest to Posttest by Intervention Group.

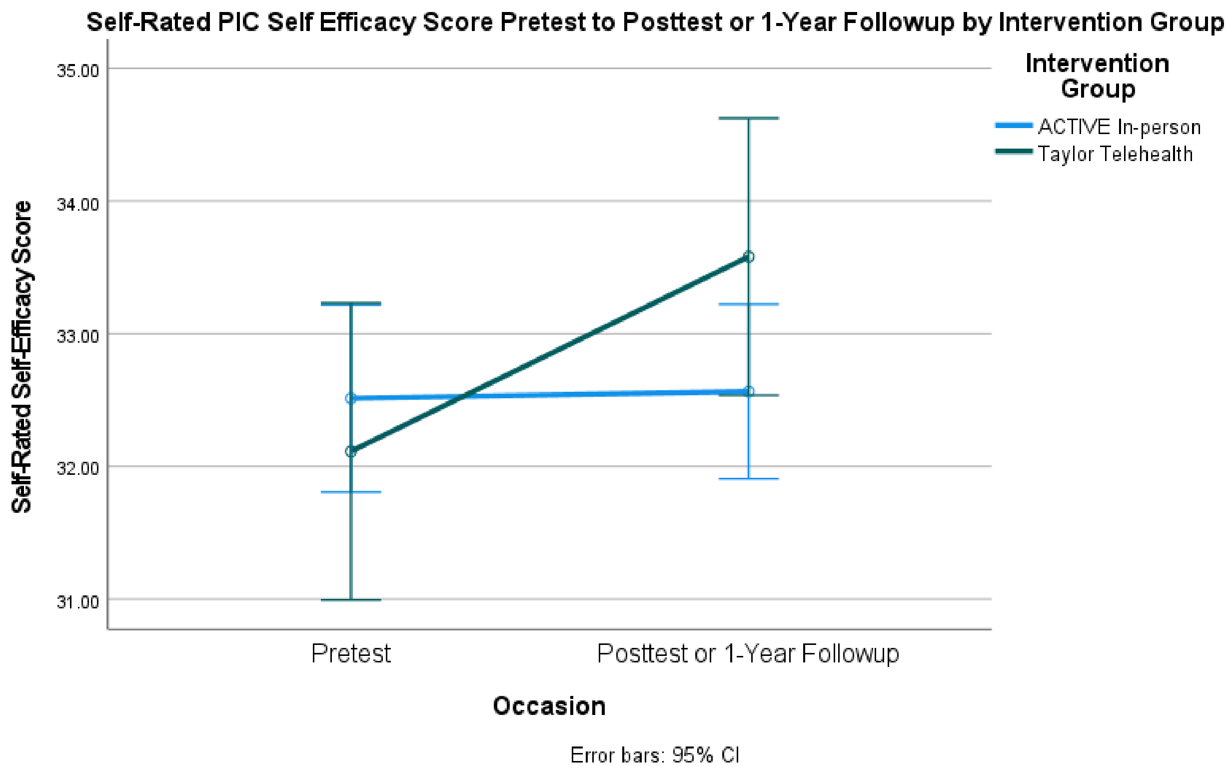


Figure 4-23. Self-Rated PIC Self Efficacy Score Pretest to Posttest or 1-Year Followup by Intervention Group.

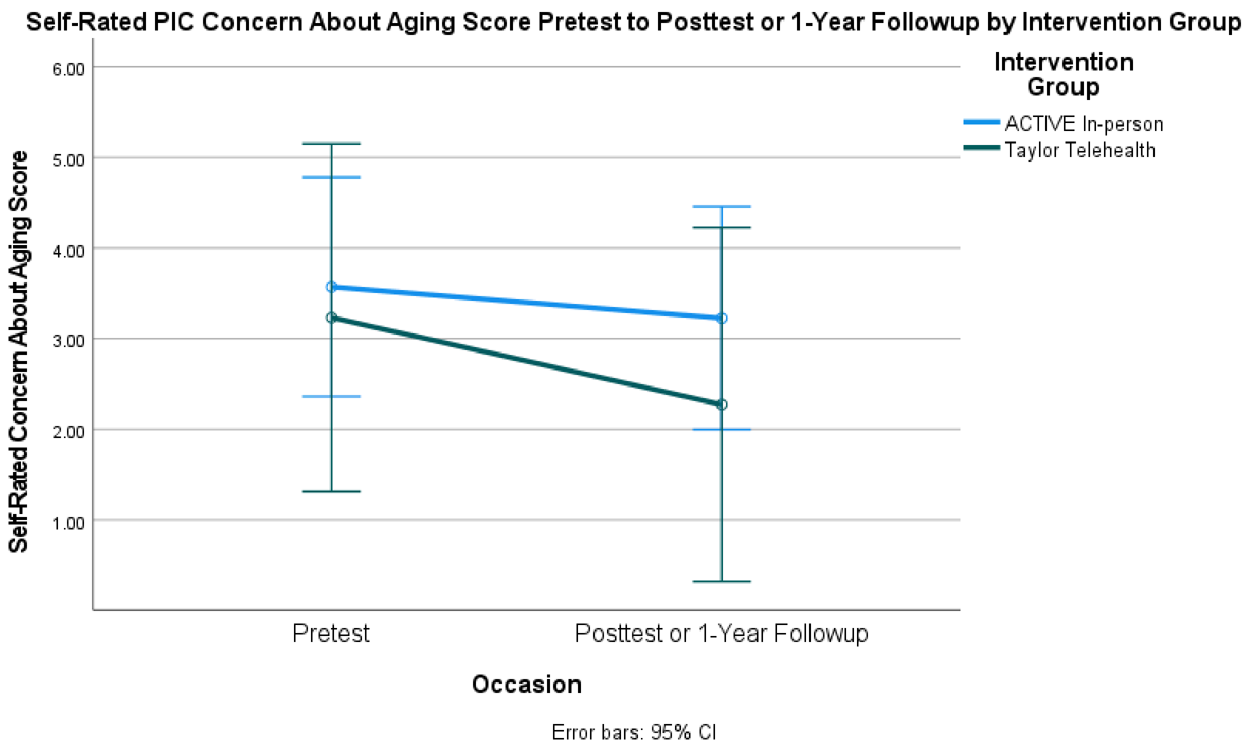


Figure 4-24. Self-Rated PIC Concern About Aging Score Pretest to Posttest or 1-Year Followup by Intervention Group.

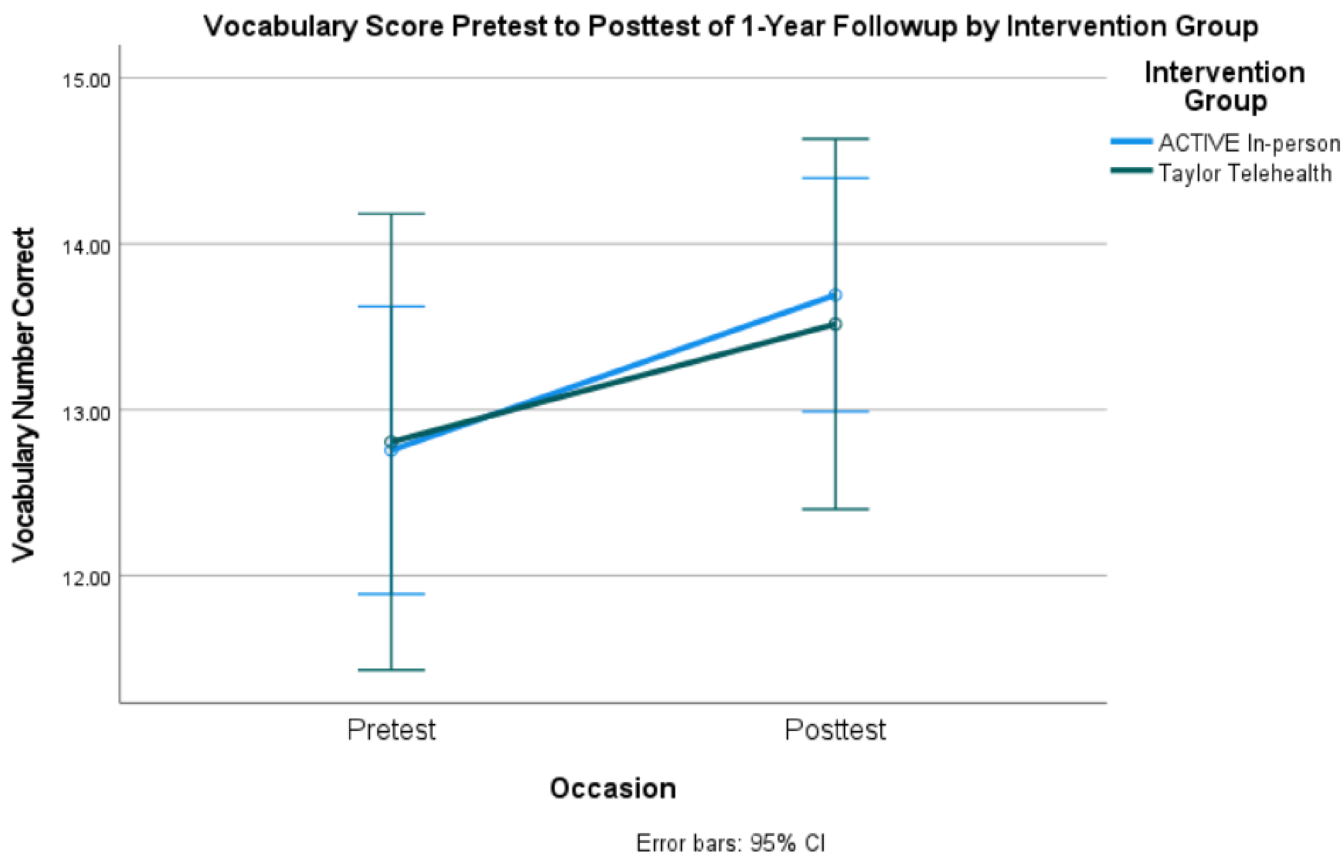


Figure 4-25. Vocabulary Score Pretest to Posttest by Intervention Group.

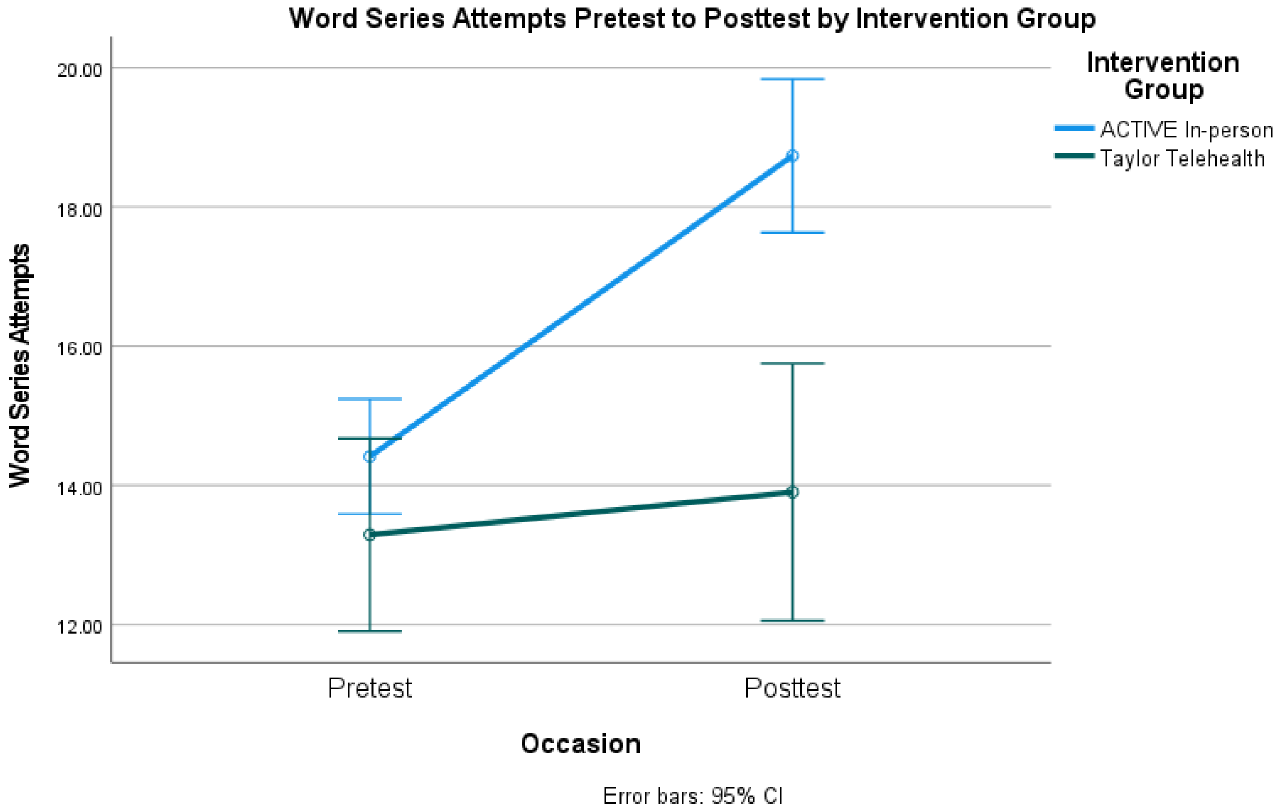


Figure 4-26. Word Series Attempts Pretest to Posttest by Intervention Group.

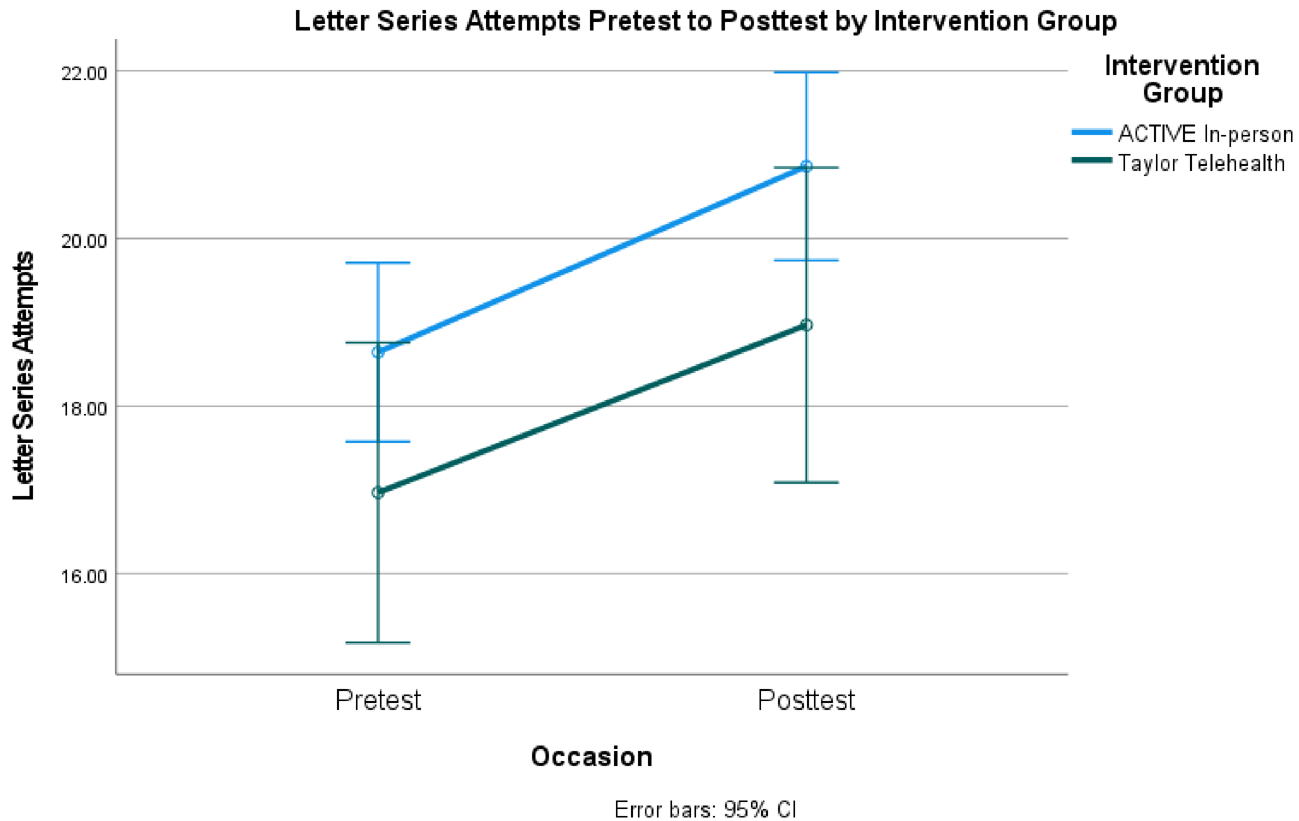


Figure 4-27. Letter Series Attempts Pretest to Posttest by Intervention Group.

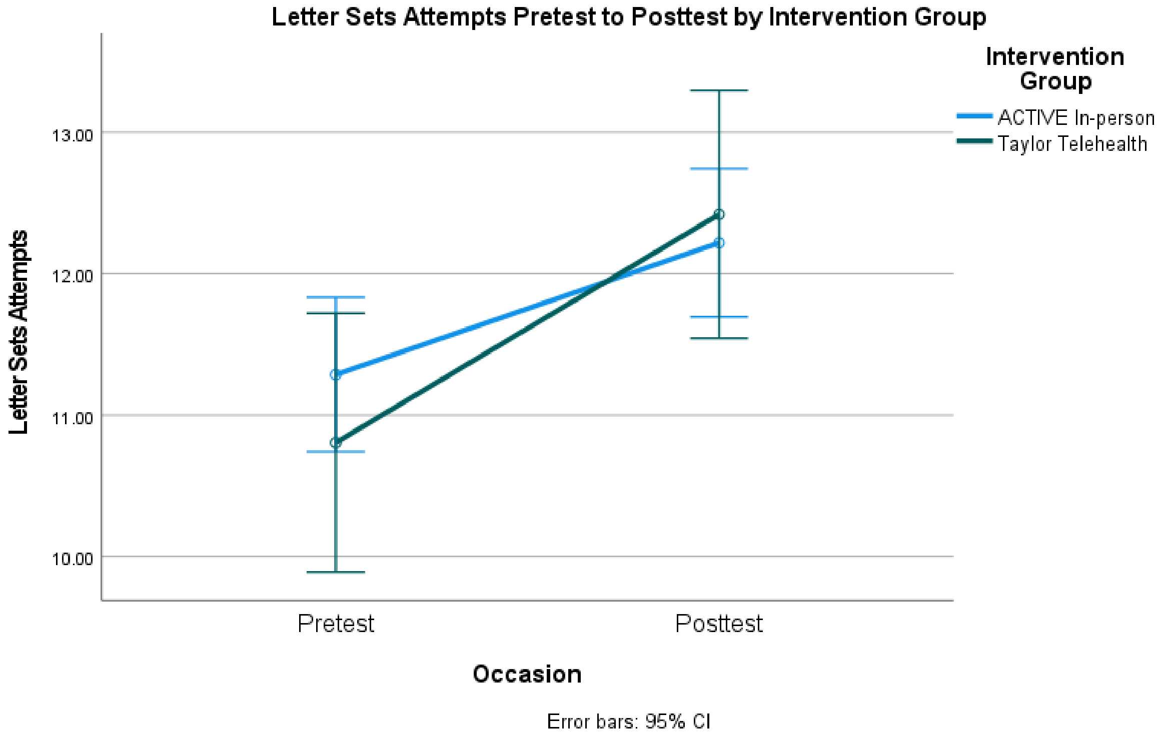


Figure 4-28. Letter Sets Attempts Pretest to Posttest by Intervention Group.

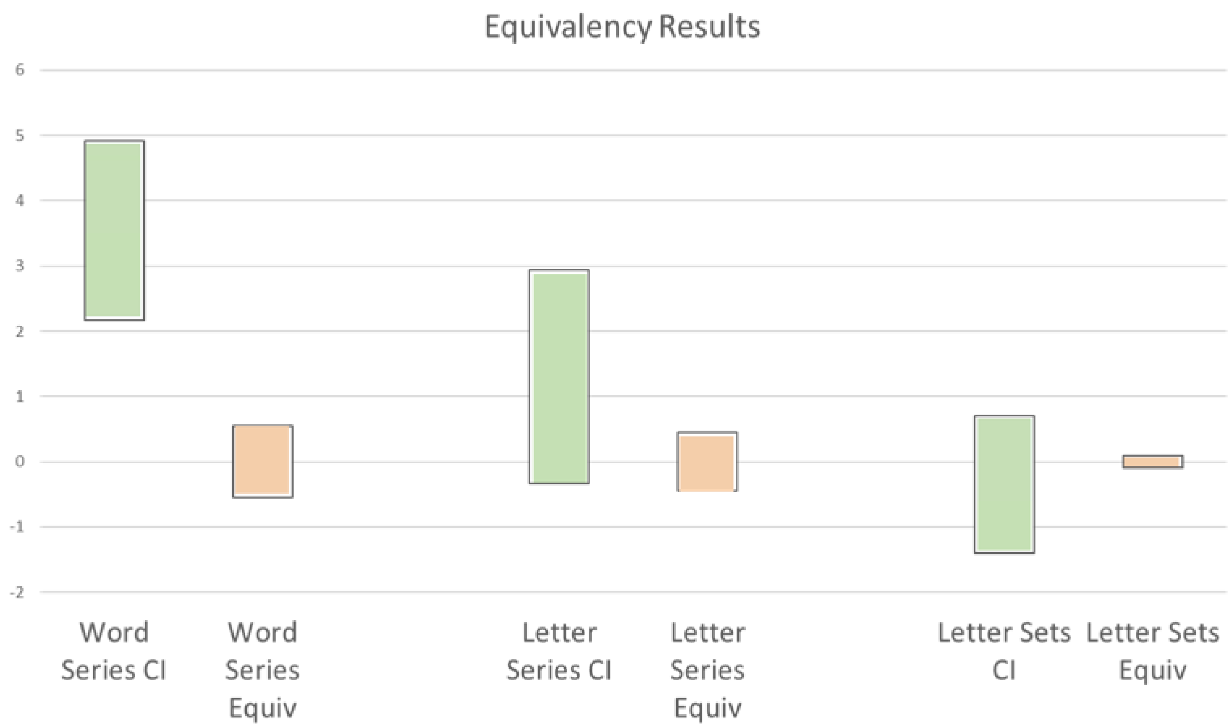


Figure 4-29. Equivalency Analysis Results. CI = confidence interval, Equiv = 10% zone of equivalence of the difference.

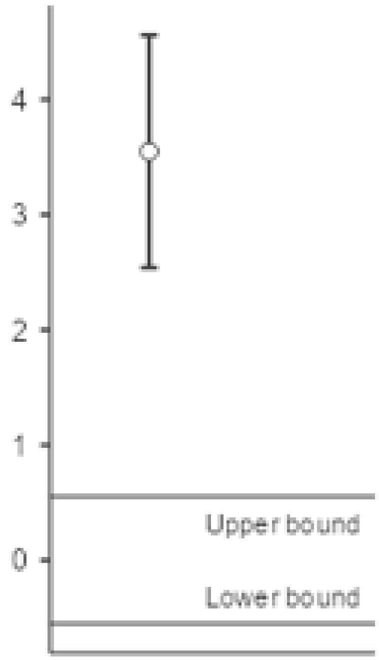


Figure 4-30. Word Series Confidence Interval and the 10% Equivalency Interval.

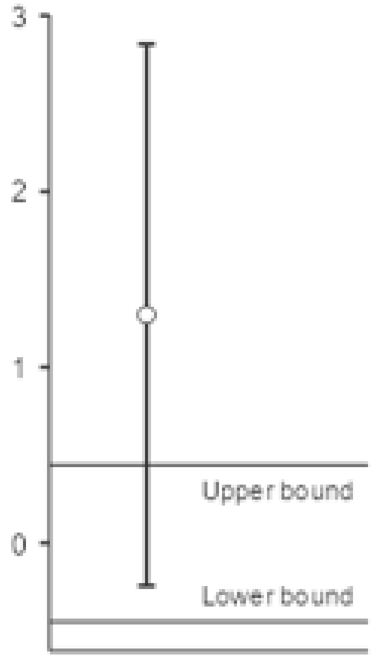


Figure 4-31. Letter Series Confidence Interval and the 10% Equivalency Interval.

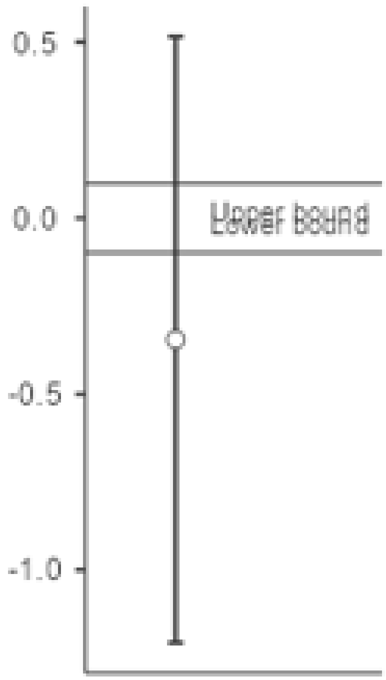


Figure 4-32. Letter Sets Confidence Interval and the 10% Equivalency Interval.

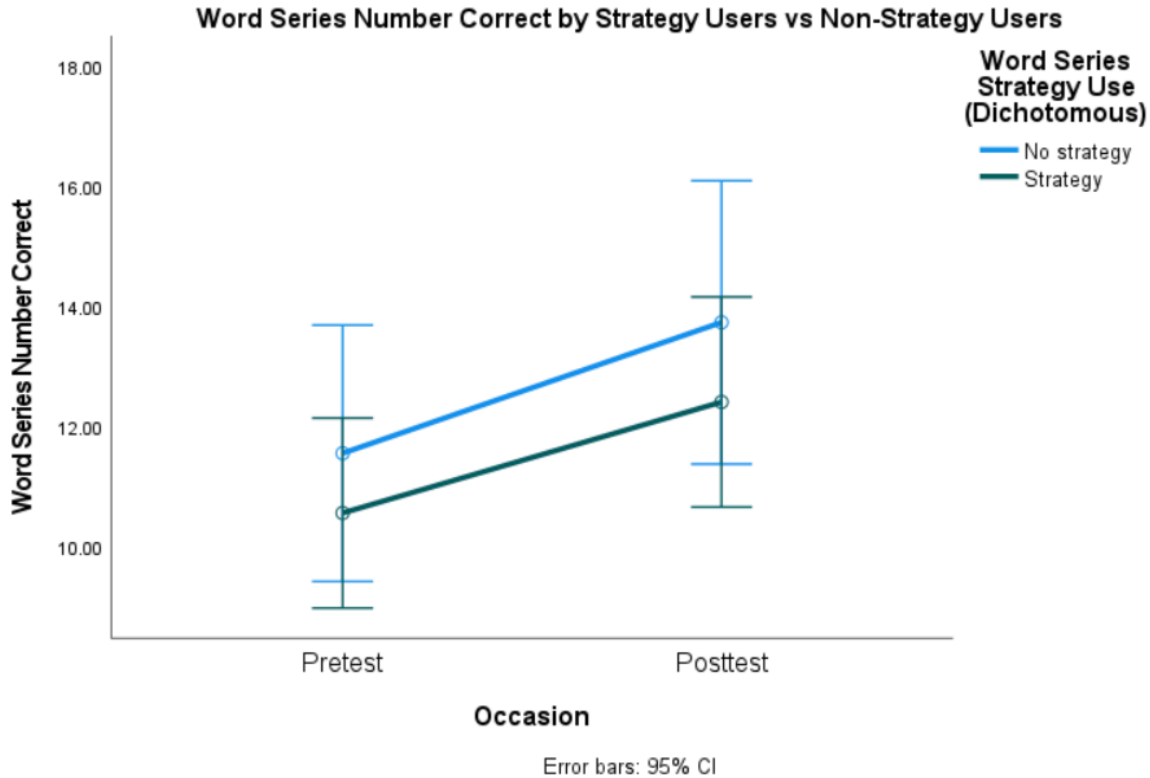


Figure 4-33. Word Series Number Correct by Strategy Users vs Non-Strategy Users.

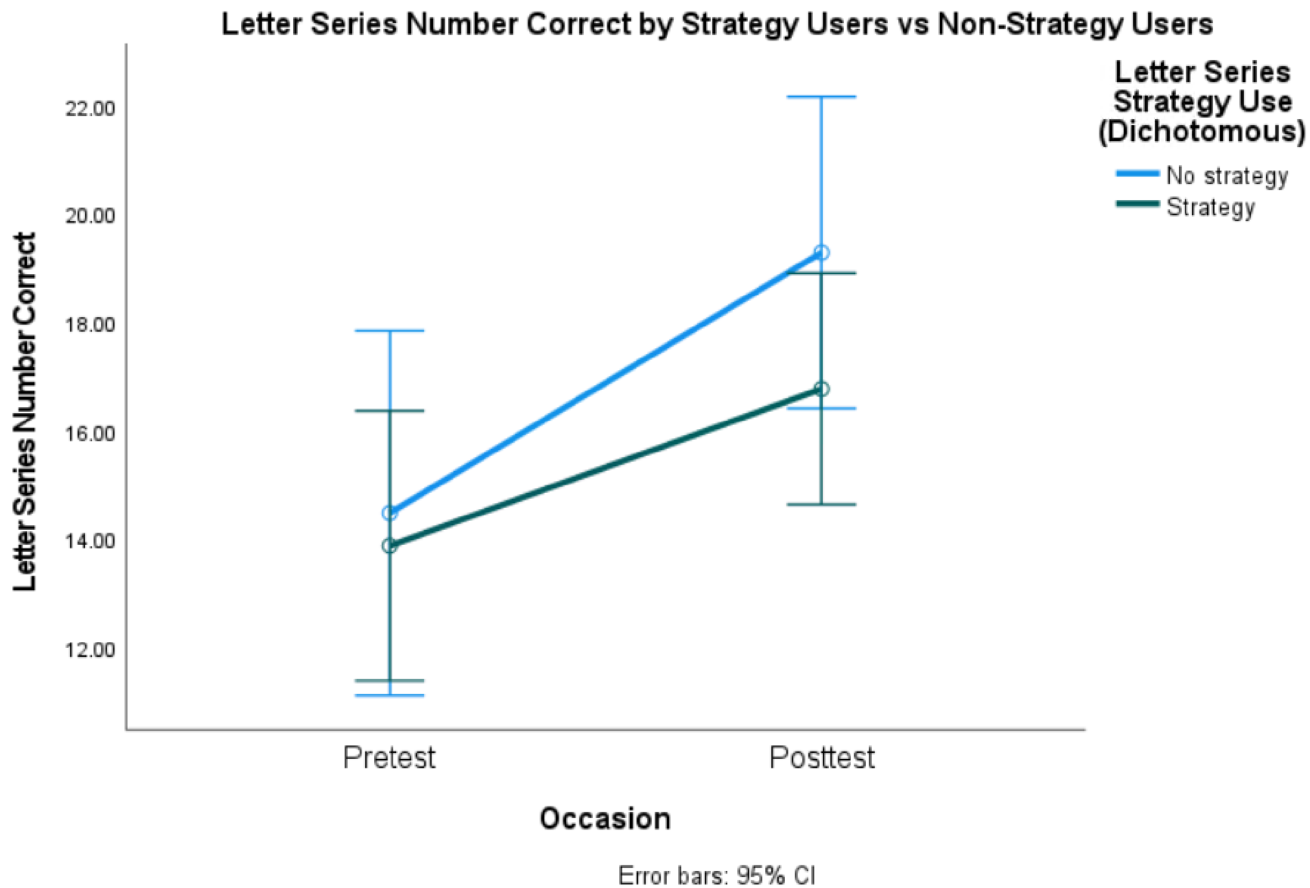


Figure 4-34. Letter Series Number Correct by Strategy Users vs Non-Strategy Users.

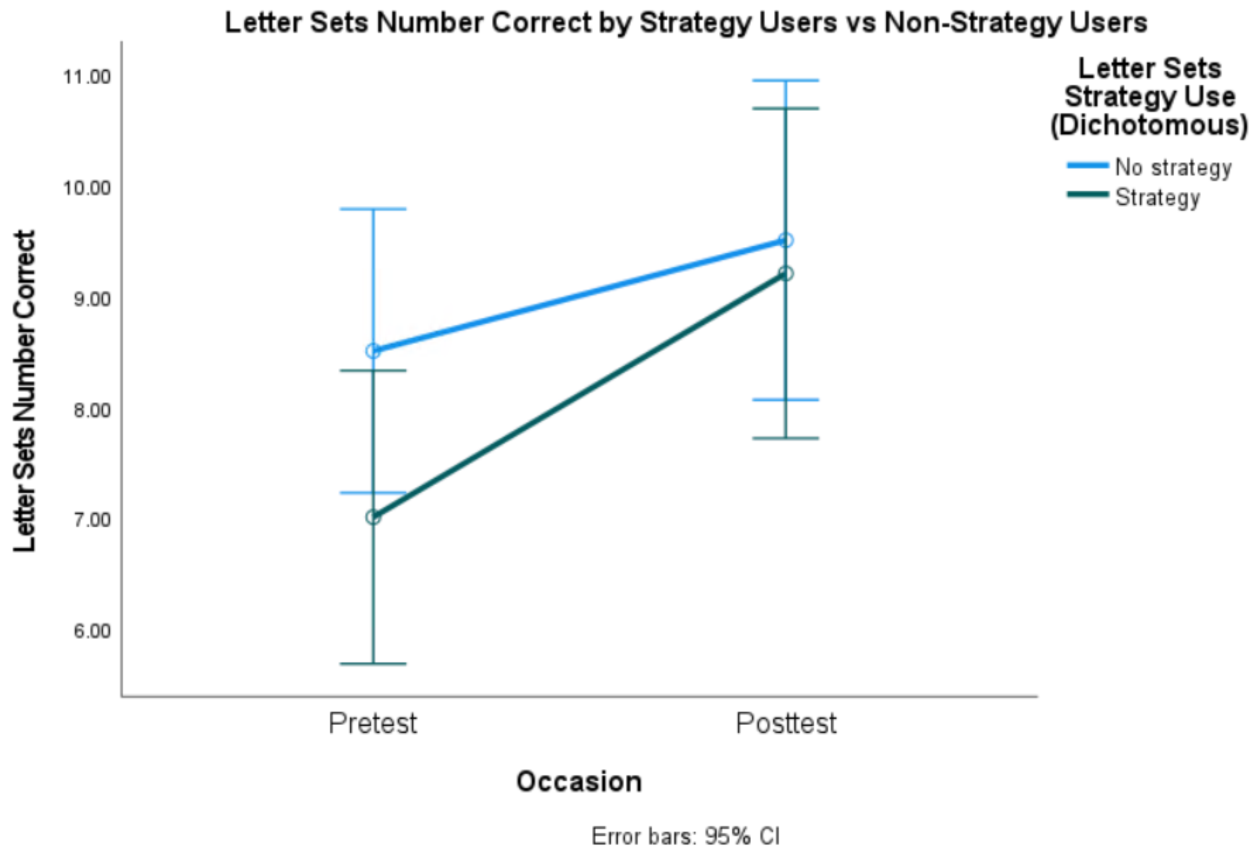


Figure 4-35. Letter Sets Number Correct by Strategy Users vs Non-Strategy Users.

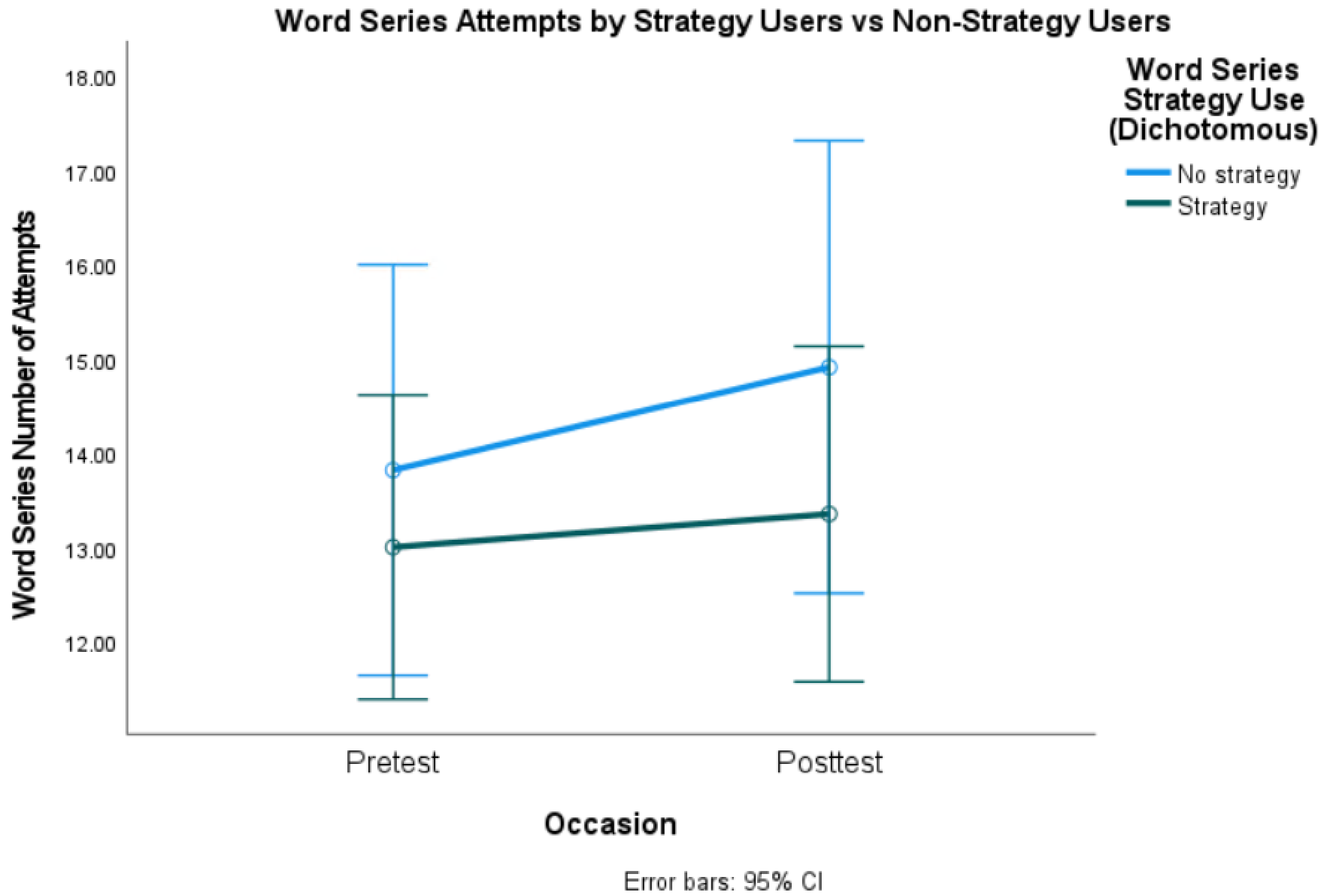


Figure 4-36. Word Series Attempts by Strategy Users vs Non-Strategy Users.

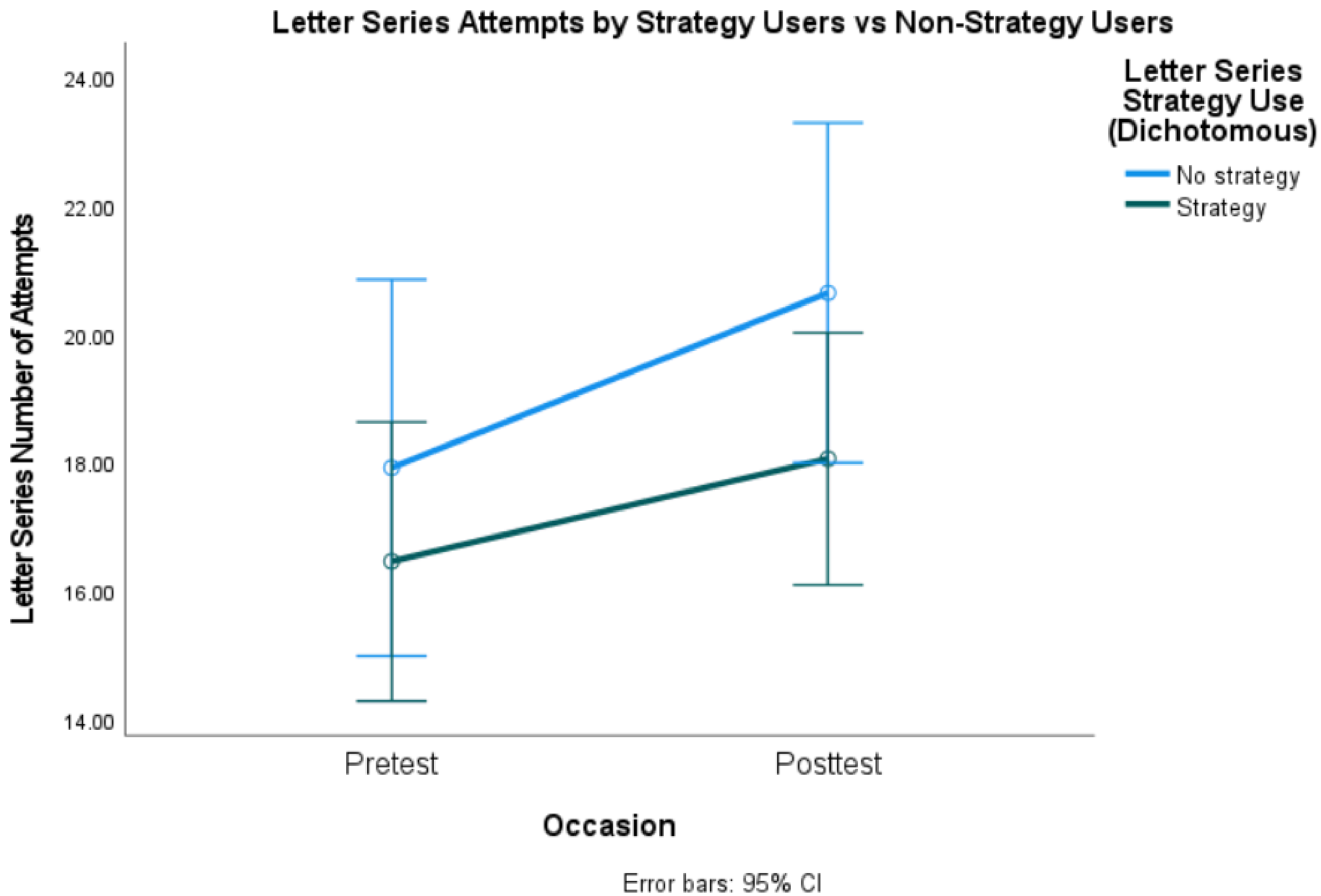


Figure 4-37. Letter Series Attempts by Strategy Users vs Non-Strategy Users.

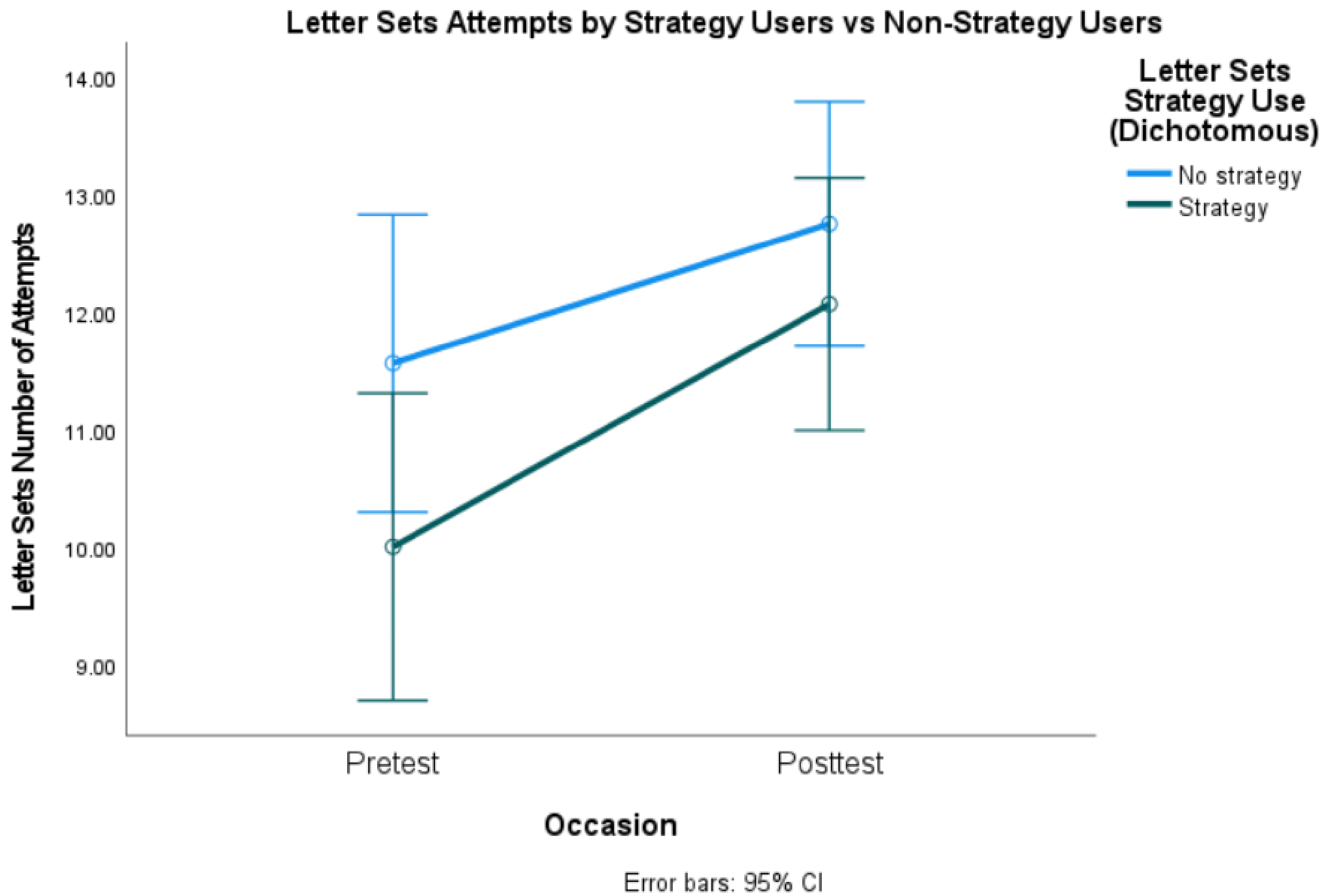


Figure 4-38. Letter Sets Attempts by Strategy Users vs Non-Strategy Users.

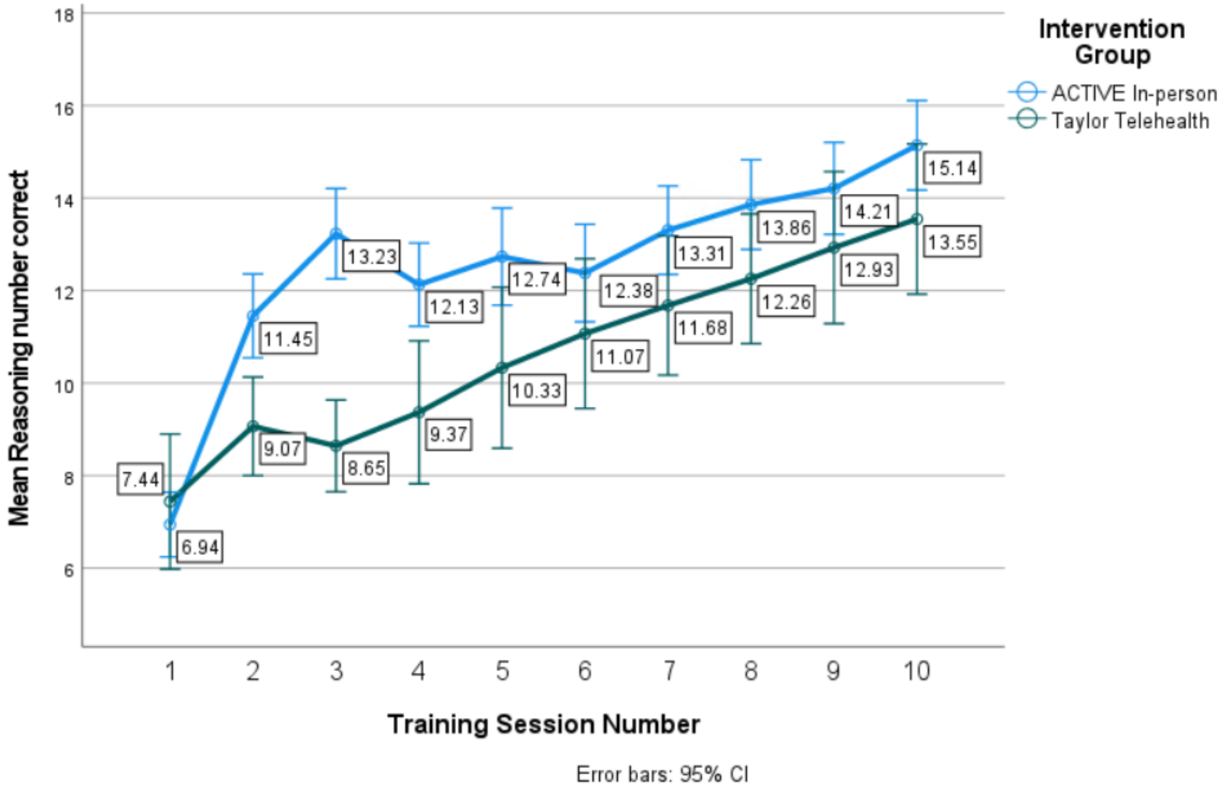


Figure 4-39. The Number of Correct Reasoning Items by Training Session Number for each Intervention Group. Note, values given in rectangles are the mean for each group at each occasion.

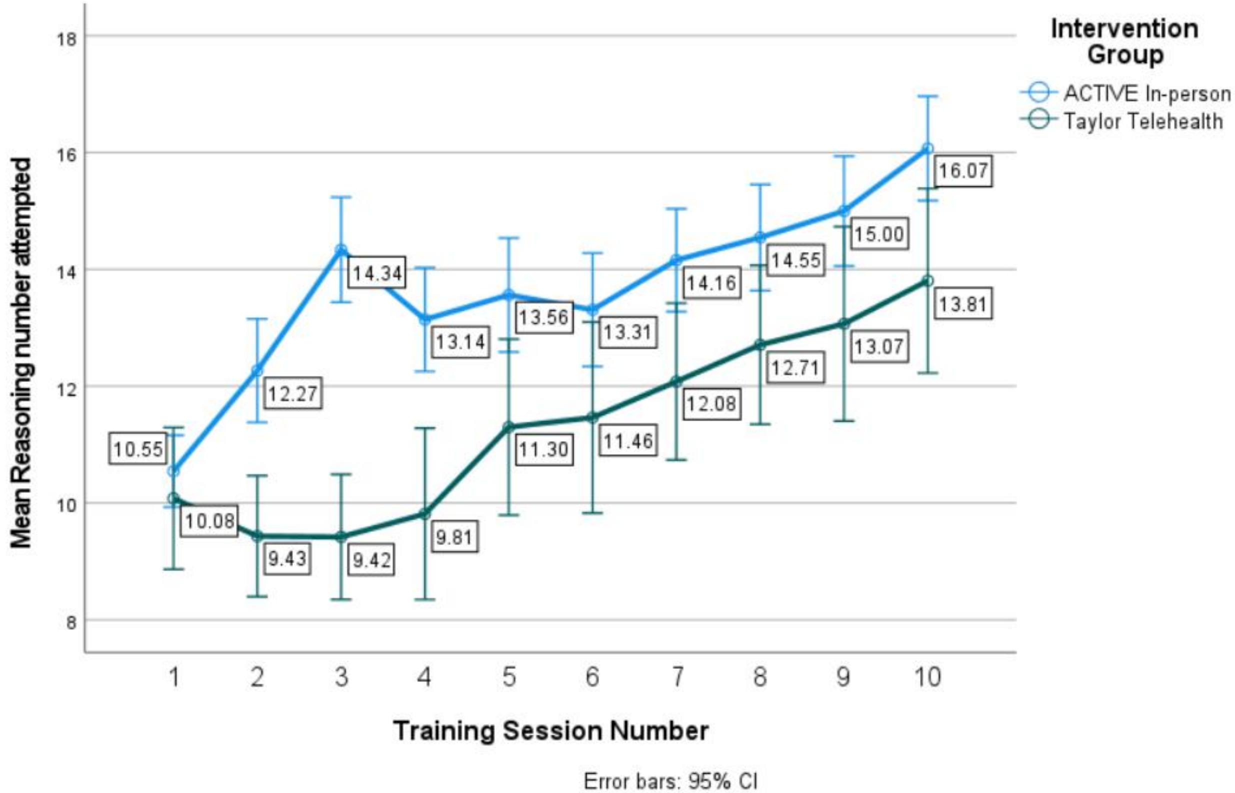


Figure 4-40. The Number of Attempted Reasoning Items by Training Session Number for each Intervention Group. Note, values given in rectangles are the mean for each group at each occasion.

CHAPTER 5
DISCUSSION

Chapter 5 is organized into five sections. First, we briefly summarize key findings of the planned and follow-up analyses. Second, we consider the conceptual implications of these findings, in light of our literature review and prior hypotheses. The third, fourth and fifth sections discuss limitations, future directions and conclusions, respectively.

Key Findings

In general, the results of the telehealth intervention were supportive of telehealth cognitive interventions. Participants rated the intervention positively, and training effects were comparable to the in-person ACTIVE inductive reasoning trial for two of the three studied primary outcomes.

Based on a post-training feasibility and useability questionnaire, most participants in the telehealth intervention found the intervention to be useful; Zoom to be easy to learn to use, simple, and enjoyable; felt they could communicate well; and were generally satisfied and would use telehealth services again. Participants were mixed in how they responded to a question asking if the visits provided over the telehealth system have the same value as in-person visits. It should be noted that participants had no in person visits, so they were speculating compared to their expectations about what face-to-face training sessions might have looked like.

The mixed evaluation of telehealth relative to (idealized) in person is informative. The main idea of this study was to modify the setting from in-person to telehealth and, while we sought to best mimic the ACTIVE in-person inductive reasoning trial, there are inherent differences that need to be considered (as discussed in the qualitative section of results). Technology challenges, different interfaces, and the need for technology assistants all added burden to telehealth that would not have been encountered in person; these were likely offset by the convenience of home visits that required less time, transportation, and parking. That the overarching impression of these evaluations was positive seems to be encouraging for the feasibility of future telehealth cognitive training interventions.

Overall, Taylor telehealth participants improved more from pretest to posttest on Letter Series and Letter Sets measures (2 of the 3 inductive reasoning tasks) than untrained control participants from the ACTIVE study; this was not true for the Word Series measure, which we consider in greater detail below. The Taylor telehealth participants also endorsed greater improvements in self efficacy and less concern about aging at immediate posttest when compared to untrained ACTIVE controls (who did not get their post-training assessment of these believes until one year after training). Our far non-Reasoning transfer measure of Vocabulary did not differ in pre-post change between Taylor telehealth and ACTIVE controls, as expected, since this was not a target of training. Follow-up analyses suggested that Taylor telehealth participants improved less on one of the inductive reasoning tasks (Word Series) from pretest to posttest. We speculate that this is due to special challenges of completing the Word Series in our chosen online administration in Canvas. Participants had to do much more scrolling and page turning in the online version than they would have had to on the original paper-and-pencil task. The motor/time demands of these extra page turns might have constrained the number of attempts (and thus the amount of improvement) that telehealth trained participants could make. However, the Taylor telehealth participants improved more on two of the inductive reasoning tasks, and for Letter Sets, this was congruent with a greater increase in the number of attempted items relative to controls. However, the association between increased number of attempts and score improvement was also not true for Letter Series, where attempt improvement was similar between the groups.

When compared to ACTIVE inductive reasoning trained participants, Taylor telehealth participants did not differ significantly in their pretest to posttest improvement on Letter Series and Letter Sets measures. Again, Word Series emerged as different; ACTIVE trained participants demonstrated significantly greater pre-post improvements than Taylor telehealth participants.

Somewhat surprisingly, Taylor telehealth participants demonstrated greater improvements in self-efficacy from pretest to posttest. However, this finding is complicated by the fact that ACTIVE used a different follow-up interval for this measure (1-year post training, rather than immediately as in the current study). Thus, we cannot exclude the possibility that ACTIVE also would have had such immediate self-efficacy improvements, but they had dissipated by the time of the one-year follow-up. Vocabulary performance change was similar between the Taylor telehealth and ACTIVE in-person training groups, which is consistent with expectations and prior literature on limited far transfer of training.

Follow-up analyses again suggested that, compared to ACTIVE participants, Taylor telehealth participants showed less increase in the number of Word Series problems attempted from pretest to posttest (while ACTIVE trained participants completed a much greater number of problems at posttest when compared to pretest). This appears to mirror the finding that the ACTIVE in-person trained participants also demonstrated a much greater increase in the number of correct problems at posttest than pretest on Word Series. Similarly, on Letter Series and Letter Sets, there was no group difference in pre to post change in the number of attempted problems, as was the case that there was no group difference in pre to post change in the number of correct responses. This suggests that the telehealth trained group evinced similar improvements in efficiency and accuracy of responses as the ACTIVE in-person trained group on these latter two measures.

Statistically, we lacked the power to claim equivalent training gains for the three inductive reasoning outcomes. Likely due in large part to a small sample size (i.e., we were underpowered for equivalency analyses by over 1,000 cases), we could only reject the null hypotheses of equivalent change between the two training groups for all 3 inductive reasoning tasks. Thus, we cannot claim that training gains were the same (on Letter Sets and Letter Series), despite the apparent similarity of improvement.

We also sought to find evidence of whether the trained strategies in the current study could be specifically tied to performance improvements. Did participants who actually *used* the trained strategies at posttest show evidence of better performance? (Alternatively, if using strategies explicitly via paper and pencil, rather than internally, slows people down, evidence of strategy use could actually be associated with poorer performance). We could record strategy use (evidence of underlines, slashes, other trained markings) only for Taylor telehealth participants; strategy data were not available in ACTIVE). A comparison of those who did and did not use strategies yielded no evidence of difference at the posttest. It is worth noting that our small sample was further divided into two groups for these analyses (strategy users vs. non-users), and this likely meant insufficient power for this exploratory analysis.

Looking within the training sessions themselves, an end-of-session test was given after each of the ten sessions. This same test was given in both ACTIVE Reasoning training and the current study. Analysis of end of session data did not look at the pretest or posttest, and thus was independent of the planned analyses. By inspection of the performance curves in both groups, there was roughly parallel improvement between the training groups across the ten training sessions. Also, by inspection, ACTIVE participants seemed to experience a substantially larger gain between first and second session, and then maintained much of this advantage for the remainder of the training sessions. This meant, at the end of the training, the ACTIVE participants were getting about two more items correct than the telehealth participants. While not significant (there was no group*time difference in the acquisition curves of the two groups), it raises the possibility that the two-point difference was a “telehealth handicap” or “telehealth differential”, representing a rate limiter on the magnitude of improvement in telehealth participants. However, we should note that this 2-point difference is small and unlikely to be clinically meaningful. But because there was no clear evidence of asymptote, it is possible that with additional sessions the telehealth participants would have caught up, but of course this would constitute an additional training burden. It should also be noted that the ACTIVE in-person training group had a similar rate of improvement in the number of attempted problems over the ten training sessions as the Taylor telehealth group, albeit again with that two-point differential (ACTIVE in-person attempted about two more problems per session).

Conceptual Considerations

In the next section, we now consider some of the larger questions that were raised in the introduction, or that emerged after consideration of the study findings.

Positive Study Design Features

We began this study by arguing that it addressed a current gap in the cognitive training literature. By examining the feasibility and effectiveness of a telehealth-delivered cognitive training intervention in inductive reasoning, this study represented (to our knowledge) the first inductive reasoning training intervention to be administered over telehealth. The current study benefited from an adequate sample size and was pre-registered at clinicaltrials.gov. Additionally, it benefited from strong adherence to the original ACTIVE in-person inductive reasoning training protocol, with the guidance and asynchronous-supervision of a Principal Investigator from the original ACTIVE trial (Michael Marsiske). While we were able to leverage the ACTIVE study for control and alternative treatment participants, the current study did not have (a) a new control condition in which participants received telehealth-administered assessments, or (b) an active control condition (e.g., alternative treatment), which the parent study had. In addition, ideally, we would have conducted at least a 1-year follow-up visit to better mirror the additional follow-up periods (after immediate posttest) that were conducted in ACTIVE trial. This would also have allowed us to administer Vocabulary and self-belief (self efficacy, concern about aging) measures in the same time frame of ACTIVE, and also to better understand longer-term durability of training via telehealth versus traditional face to face.

Why was Word Series so different?

The lack of discernible training benefit on Word Series via telehealth was unexpected. (Letter Series and Letter Sets behaved as expected: improvement was significantly better than in controls and not significantly different from ACTIVE Reasoning participants). As noted in our qualitative findings in the Results section, we suspect that this is an interaction between the presentation of Word Series items (presented vertically down a page, rather than horizontally across the page as in the other reasoning measures) and the technology. Specifically, the Canvas interface and screen sizes meant that many more page turns were needed to complete all Word Series items in the telehealth condition than in person. Not only could this have slowed individuals (motor demands of page turns), but it could also have broken participant flow and served as a distracting stimulus. It seems unlikely that anything inherent in Word Series responded differently to telehealth, because the series rules used are *identical* to the series rules used in Letter Series (which showed improvement). Word Series was also always the first Reasoning measure administered (which was not always the case in ACTIVE), and thus may have benefited from less "warmup" in Taylor telehealth than in ACTIVE.

Telehealth training improved response efficiency

When examining the number of items attempted by the participants, telehealth-trained participants increased more in their number of attempts from pretest to posttest than untrained ACTIVE controls on Letter Sets, but ACTIVE controls increased more in their number of Word Series attempts, and there was no difference in Letter Series change. In addition, the ACTIVE inductive reasoning trained group demonstrated greater improvements in the number of attempted Word Series problems from pre to posttest than the telehealth trained participants, but there was no difference between the groups in pre-post change in Letter Series and Letter Sets attempts,

Considering the same pattern of findings were found in pre-post change in the number of correct responses between the two training groups, the results suggest that the two interventions act similarly in the way improvements were achieved, such that both interventions seem to improve accuracy and efficiency to a comparable degree. Neither intervention evinced greater accuracy improvements (i.e., where the proportion of attempted items correctly responded to increased) =or greater efficiency improvements (less time needed per item, resulting in more attempted items) than the other. Overall, our findings seem to suggest that for two of our three reasoning measures, the magnitude and nature of training gains were similar when telehealth was compared to in-person ACTIVE training. Speculatively, as noted above, the lack of a comparable effect for Word Series is likely due to the unique motor and time demands that the online version of Word Series (used in telehealth) imposed over and above the paper-and-pencil version (used in in-person training).

Are telehealth-associated improvements in self efficacy real?

Those in the telehealth-trained group self-reported greater improvement in self-efficacy when compared to the ACTIVE no-contact controls and ACTIVE in-person trained participants. The important caution to make here is that the current study assessed self efficacy and concern about aging immediately before and after training, whereas in ACTIVE the follow-up assessment of these self beliefs did not occur until one year after training. Thus, it is possible ACTIVE participants who were trained would have evinced similar improvements in an immediate posttest. That said, given the anecdotal reports of greater ease and comfort with computer use, and the increase in the number of reasoning items attempted, one might also speculate that self-efficacy improvements were linked less to training, and more to their burgeoning comfort with technology. ACTIVE in-person training was not associated with self-efficacy change (Parisi et al., 2017), and this is consistent with findings from other training interventions. For example, in HABIT, those who received Memory Support System (calendar) demonstrated memory self-efficacy improvements (Greenaway et al., 2013); however, those who received Computerized Cognitive Training did not demonstrate any memory self-efficacy improvements (Chandler et al., 2019). This may suggest that training studies involving social interaction and compensatory technology (calendars) are associated with self-efficacy improvements, even though direct in-person cognitive training may not be. In the telehealth-trained inductive reasoning sample, it is likely that the social interaction and, perhaps more uniquely, the introduction to new technology could have conferred the self-efficacy benefits.

Implications of the useability and feasibility assessments

Overall, the participants found the training to be useful, easy to use, easy to interact with, easy to communicate about, and they were satisfied with the training program. This, along with the low incidence of dropout, is encouraging for pursuing future telehealth mediated training. That said, the sample recruited for this study was clearly in the young-old age range, was racially homogeneous (most were White), and were highly educated. The study therefore has little to say about the feasibility of telehealth delivered interventions in less advantaged populations. It is easy to imagine that older adults from lower SES groups, for example, might be less likely to have a computer and mouse. In this study, persons who only had more widely available smart phones and tablets (which are less available) could not be enrolled, due to the inability to reliably use the Canvas e-learning platform. Lower SES groups might also have less experience with computers and might also start at initially lower levels of reasoning. All these factors could represent substantial threats to scaling the current intervention to a more heterogeneous population. Thus, more data from a more broadly representative population of older adults is needed.

Limitations

No active control

The present study was limited by the lack of an active control group. A no-contact control group was culled from the original, in-person ACTIVE trial. This group received pre-test and post-test cognitive testing and questionnaires but did not receive any intermediate intervention of any form. ACTIVE itself had "active" control groups, because there were three alternative interventions (reasoning, memory, speed) that could be compared against each other. In the current study, having an active control group that also was assessed with remote protocols (Zoom, Canvas) would have produced a group that was matched in (a) generational position, (b) technology demands of assessment, and (c) expectations and engagement (if an alternative or placebo intervention had been offered). Having an active control group, or one in which the participants complete another activity in place of the cognitive training intervention, allows for an experience more comparable to that of the cognitive training intervention (Simons et al., 2016). An active control group can reduce motivational differences between the training and control groups and help equate the social contact. Because the current study lacked an active control group that was matched in technology demands to the intervention group, we cannot rule out motivational and social contact differences as possible confounding variables that helped explain differences between our telehealth training groups and the ACTIVE no-contact control group. Having a Zoom-only placebo control group would be one way to separate the non-specific training effects (how to use a device) from inductive reasoning-specific training results in future research.

Altered certification in ACTIVE training instruction

ACTIVE had a protocol for being certified as a trainer, which involved reviewing standard study procedures, followed by a demonstration of each session. Then, the trainer practiced with a mock participant, followed by an older adult, non-participant volunteer. These were video-recorded and reviewed by an already certified trainer to give session-to-session feedback. ACTIVE's manual of operations laid our certification procedures (https://osf.io/nyad9/).

The present study's trainer, Brad Taylor, reviewed the standard study procedures and practiced the training sessions with a certified trainer, Dr. Michael Marsiske, in addition to Research Assistants (RAs) serving as mock participants, along with an older adult volunteer. In addition, once in the field, two real training sessions were recorded, with participants' permission, and were subsequently reviewed by Dr. Marsiske using an ACTIVE fidelity checklist. Marsiske then met with Taylor to provide feedback and suggestions for improvement to guide future sessions. Additionally, there was one session (a training session 9) that was administered by one research assistant, Ryan

Faulkner, with the proctoring assistance of another assistant, Madison Verdone. These two assistants had been heavily involved in the modification of ACTIVE's inductive reasoning training workbooks to a Canvas format, assisting with and observing many training sessions, practiced leading portions of sessions under the live supervision of Brad, provided scripts to aid with standardization, and had practiced and met with Brad multiple times before being allowed to run the session independently.

While the procedures in the current study were rigorous, certification procedures did not exactly match ACTIVE, in part because much of the ACTIVE certification was done centrally (by a coordinating center in Massachusetts; this option was no longer available). It is unknown whether differences in training and fidelity might have produced the differences observed (especially on Word Series). The process of completely matching original certification procedures would have likely conferred greater matching of trainer competency, ability to pace and organize sessions, and ability to manage the group learning environment.

Non-equivalent MMSE administration

As noted in the qualitative remarks, pretesting in the current study was done via Zoom and Canvas. Thus, the MMSE (Folstein et al., 1975), which was originally designed for in-person, paper-pencil testing, needed to be modified for telehealth delivery. MMSE screening followed established telehealth administration protocols (Cullum et al., 2014; Grosch et al., 2015). All MMSE items were the same as the original MMSE. However, participants may have been able to answer the orientation questions (What is the date? What time is it?) with the clock and calendar that are commonly displayed on most computers. However, the participants performed quite well on the

MMSE (mean = 28.58, SD = 1.31, range 26-30). Even if every participant had gained two additional points because of use of the computer display, the performance range would still have been 24-28. Eligibility the current study required an MMSE of 23 or higher; thus, regardless of this potential testing error, all participants would have remained eligible.

Training inconsistencies with ACTIVE

There are some inconsistencies in how the telehealth training was conducted, compared to the original ACTIVE inductive reasoning training. For example, while the original ACTIVE inductive reasoning training groups were a maximum of 4-5 participants, the telehealth training occasionally had training sessions with 6 participants to allow for participants' schedule flexibility and better adherence. Additionally, in the original ACTIVE trial, participants had at least a day separating their training sessions from each other, but again, to allow for participants' schedule flexibility and better adherence, some participants did make-up sessions on back-to-back days or even two sessions back-to-back in the same day (this also occurred, infrequently, with makeup sessions in the original ACTIVE study). These flexible schedulings were particularly helpful when participants anticipated time away from home and wanted to do some sessions in advance or as a make-up. Further, while the original ACTIVE trial conducted training during business hours on weekdays only, the telehealth trained participants were provided more flexibility to have training sessions in the evenings or on weekends if they wished. As mentioned in Chapter 3, in the telehealth training intervention, 10 standalone make-up sessions were conducted for 8 participants in total, and there were 35 instances of participants moving from one session to another. While the exact number of make-ups and instances of participants moving from one session to another in the ACTIVE trial is not available, this likely represents more flexibility than the parent ACTIVE study allowed. In addition, because of anticipated travel, three participants completed posttesting the day following their final training session, which was sooner than the 1-2 week window sought by the ACTIVE trial.

The use of Canvas and Zoom proved to be challenging to manage simultaneously for many of the participants, especially in the earlier training sessions, and this led to longer initial training sessions. ACTIVE's training sessions were 60-75 minutes per session, and the timing of each session was broken down with specific times allotted to each activity within the training sessions. We attempted to mimic this timing in the telehealth training, and the activity timing was adhered to in later training sessions. In the first 1-2 training sessions, however, many participants required additional time for help logging into Canvas, switching between Zoom and Canvas, and for general troubleshooting problems. Thus, on average, the first training session in the current study took approximately 120 minutes to complete.

Poorly-worded questions

Additionally, on the Telehealth Usability Questionnaire, there was one question that asked participants whether "the system is able to do everything I would want it to be able to do", and in retrospect, we found this wording problematic. It is not clear, for example, whether participants understood this to be an evaluation of Canvas, or of the broader study context. Additionally, there was another question asking whether the visits provided over the telehealth system were "the same" as in-person visits. This question was vague about what specifically was being referred to as the telehealth system (e.g., Zoom or Canvas or REDCap or the online registration system?) and in which ways the two intervention mediums might be expected to be "the same".

Future Technology Considerations

In the next section, we consider some of the possible future technology considerations for optimizing telehealth-delivered cognitive training to older adults.

One future technology modification would involve finding better learning management systems for training. We used the widely deployed e-learning system, Canvas, which was designed for academic courses and is used by schools and universities. As such, Canvas was not optimized for the kind of reliability and precision of timing that would optimally suit most cognitive training. For example, because the original ACTIVE training was paper and pencil, participants had to mark up their training workbooks. Although we could imagine ways of implementing that within the elearning technology, these would differ substantially from the ACTIVE markup rules. Thus, we chose to use a Canvas feature called Annotated Assignments. In this feature, participants use a virtual pencil to physically "mark up" a sheet of paper. This was basically a direct analogue of the ACTIVE procedure. However, recording participant markups reliably required steady internet connectivity. Any interruption in connectivity (an in-home wifi glitch, network congestion, temporary outages on the Canvas servers) would both compromise the markup activity and cause subsequent data loss (i.e., where markups made during the glitch were not retained). If future training were to maintain the manual markup procedure, a more reliable interface or program would be needed. In fact, there has been progress in designing learning management systems specifically for older adults, as in the virtual classroom of the PRISM study (Czaja et al., 2021). A future study would benefit from a more intuitive, simplistic interface that is optimized for cognitive training for older adults, and a virtual classroom might lend itself well to virtual cognitive training.

Further, the annotated assignments' feature also forces participants to engage in active saving of their work (e.g.., click a check mark to save annotations). Multiple annotations (underlines, slashes, tick marks, or circles) can be made before the check mark is selected, and all those annotations will be saved. However, this is page specific, such that one must click the check mark at least once per page for the page's annotations to be saved. Therefore, one cannot complete an entire, multi-page document's worth of annotations and then click one check mark to "lock in" all the answers. Further, this check mark was programmed such that it is immediately adjacent to the "trash can" icon, which deletes all their annotations up until the last time the checkmark was selected. These interface features meant that (a) participants had to rely on prospective memory to remember to save, and (b) manual coordination had to be precise, or work could be accidentally trashed. (Use of Trash did not automatically delete work, fortunately, because upon selecting the trashcan icon, participants would then be asked if they want to delete their annotations by clicking "OK" or if they want to cancel the deletion process by clicking "Cancel"). The trashcan icon's proximity to the checkmark meant participants would occasionally select the trashcan icon instead of the checkmark, and then would have to navigate to the "cancel" button, which resulted in unnecessary cursor travel and time taken. Ultimately, this check mark feature slowed participants down, as they had to further move their cursor and click the check mark frequently throughout the activities. Most detrimentally, this likely helped prevent some of the slower participants from having the time to finish all the problems.

A second future modification would be to extend the duration of training sessions. To help allow for slower participants to complete the training, given the slower nature of using an online training medium, compared to a paper and pencil format that would be more familiar to many older adults, and to allow for slower computers and internet speeds, additional time to complete activities would be beneficial. The downside is that many of the faster participants, and those with faster computers and internet speeds, would have more downtime while waiting for others to catch up.

A third future modification would be to standardize the training interface. For example, restricting the training program to participants with computer mice would also help standardize the training program and allow for most participants to navigate more quickly through the exercises. In the current study, for example, by allowing participants to participate with a variety of pointing devices (e.g., trackpads, mice), some trackpad participants were slower than those with mice. Data were not collected regarding whether participants were using a mouse or trackpad, and that would have been useful for predicting whether that affected training gains. Alternatively, it may be that training design needs to be adapted to shift the locus of training to tablets.

While Canvas officially allows for tablets and smartphones, in the current study the Canvas platform did not allow for tablets. iPads and android tablets were tested prior to study onset. To utilize the annotated assignments' feature (the feature where most of the activities were completed), only a mouse or trackpad offered the level of digital precision needed for markup. Computers also typically offered larger screens to display more items/stimuli. However, if Canvas were changed to allow for tablets (perhaps by incorporating styli or electronic pencils), or if a different training platform was used, tablets might possibly have been more user friendly, intuitive, and closer in design to a paper-and-pencil training.

It is worth noting that we had considered substantially modifying the training to optimize it for a computer environment. For example, Letter Series and Letter Sets could easily have been offered as online quizzes, where participants click or select an answer from a multiple-choice quiz tool. (This also would have allowed auto-scoring of items). However, this would have so changed the response format from the original ACTIVE study (and it would have required us to develop new methods of markup from the original ACTIVE strategies of underlines, slashes, and other highlighting rules) that we were concerned that this would have harmed comparability between our study and ACTIVE.

It seemed in our study that those who used strategies, but did not need to, felt that the strategies slowed them down. This was in contrast to participants who did not so easily internalize the rules of the reasoning tasks felt that explicit use of external strategies were of benefit to them. Future research with more substantial power and sample size should investigate whether explicit strategy instruction confers greater benefit to participants whose initial performance is poorer, in contrast to those who start out with better performance on the inductive reasoning measures.

Conclusion

In conclusion, a telehealth-delivered cognitive training intervention in inductive reasoning is useable and feasible. Most of the participants rated the telehealth intervention highly. Additionally, this telehealth-delivered cognitive training intervention in inductive reasoning produced improvements that seem comparable to in-person training for at least two of the three inductive reasoning measures.

Further, there is some evidence that the nature of improvements are similar between telehealth-trained participants and in-person trained participants from ACTIVE, such that changes in the number of attempted problems correspond very well to changes in the number of correct responses. However, the telehealth training group's acquisition over the course of training seemed to be slightly more gradual than in-person training, which led to a small net deficit in the amount of skill acquisition after ten sessions. This could be partly explained by the concurrent demands imposed, especially initially, by having to understand how to navigate the Canvas system. As the training sessions went on, study staff anecdotally observed that participants were faster to switch between Canvas activities, select the right tool (paintbrush icon or "T button" for typing answers), and draw annotations and circle the correct answer choices. Overall though, it still seems fair to conclude that telehealth training produced similar gains in inductive reasoning as ACTIVE in-person training on two of three measures studied. These findings support the idea that participants acquired explicit reasoning skills, above and beyond any more general technological benefits conferred by engagement with telehealth training.

The academic e-learning interface (Canvas) was not optimal, and perhaps the training design itself was not optimal for many online older adult learners. Canvas was susceptible to crashes and the annotated assignments feature frequently struggled to load, resulting in frustration among many of the participants. The design choice to include the check mark to certify one’s annotations was required for every annotation completed and was placed inconveniently adjacent to the trash can icon to delete one’s annotation, which resulted in occasional accidental annotation deletions. Additionally, when talking with prospective participants, many reported owning a tablet and smartphone, but not a laptop or desktop computer. Canvas’ annotated assignments features curiously does not work with tablets and smartphones, so a future telehealth cognitive training study may seek an e-learning platform that not only provides a simple, user-friendly interface for older adults, but one that also allows for tablets and smartphones.

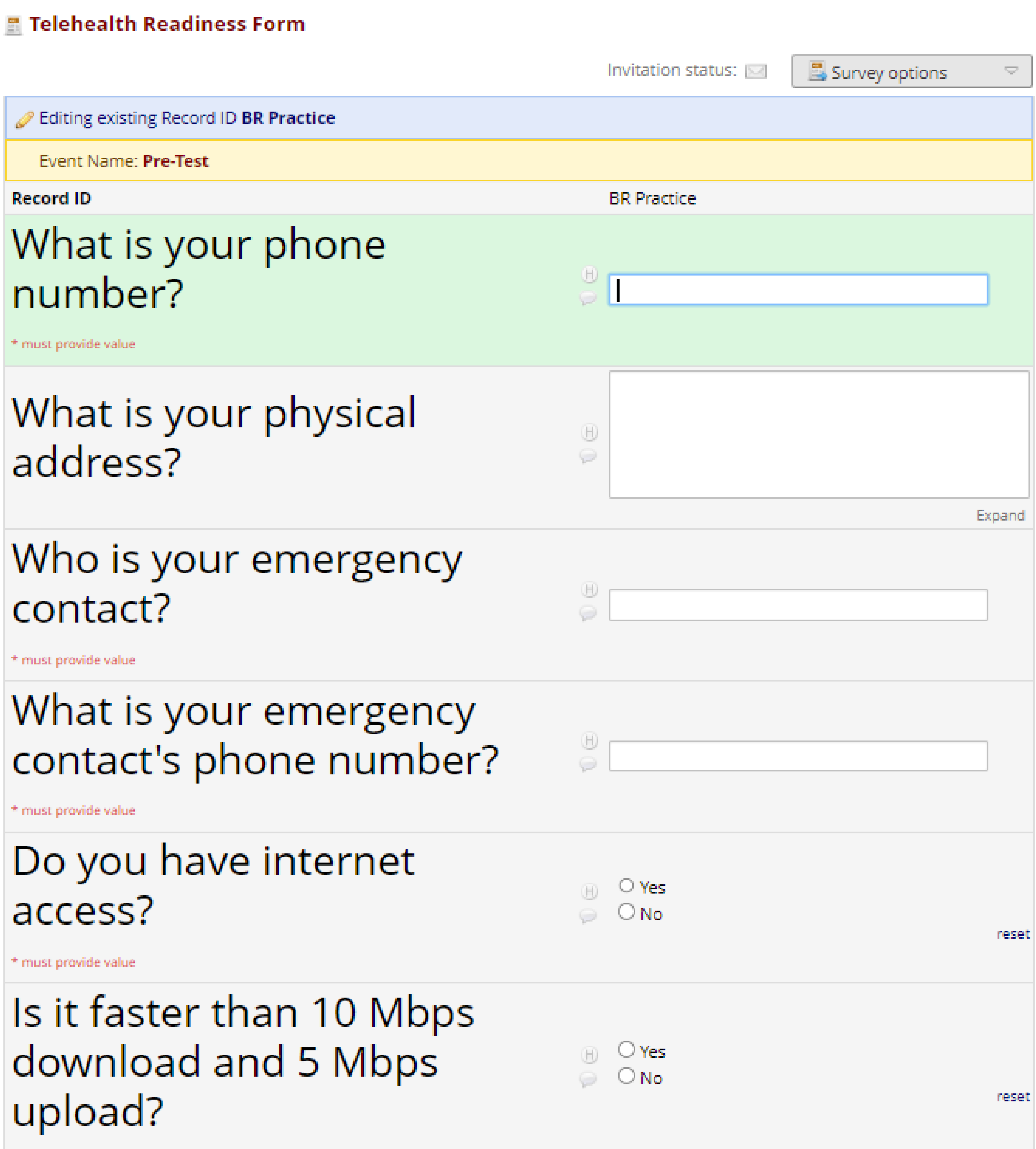
By conducting a telehealth study, only those with access to adequate technology and enough technical experience to operate Zoom and Canvas were able to participate. Due to inadequate funds to afford access to technology and IT support staff or technology trainers, individuals without adequate equipment or technological experience (desktop or laptop computer, computer capable of Zoom, experience using a computer etc. . .) could not participate. In fact, 10 (12.7%) of the 79 participants who contacted the study staff about possibly participating in this study were unable to participate because they did not meet the technology requirements. A future study may need to budget for standardized training equipment to eliminate this as a source of future individual differences. Of those who were eligible and did participate, many had different training experiences due to differences in technology, including slower internet speeds, using a computer track pad – as opposed to a mouse, or having a slower computer. Over time, the interventionists were able to learn whose computers took longer to load and were able to provide adequate time for the slowest participants to catch up when transitioning from activity to activity. However, this meant some participants felt they were slowing the group down, while other participants were forced to wait patiently for other participants' computers to load. Importantly, this also affected how easily and quickly participants could navigate and scroll through Canvas activities. Participants with faster computers, faster internet, and computer mice could often complete more activities than participants with slow computers, slow internet, and trackpads. Further, given the very-skewed, positively selected nature of this study, a future study needs to also budget adequately for recruitment.

A future training design may want to allow for more self-paced and computer administered instruction (like BrainHQ by PositScience). This deals with the limitations of needing a trainer and coordinating schedules. This would have the potential to also make an intervention widely deployable. However, this is also a radical redesign, and thus would need to be carefully tested.

Additionally, the current study failed to examine long-term maintenance, or outcomes beyond the primary target of intervention. For example, everyday cognition and wellbeing are important outcomes that ACTIVE examined for years following training, but that the scope of the current study precluded. Future trials should seek to examine whether long-term maintenance is comparable to in-person trials and examine outcomes beyond inductive reasoning.

Inductive Reasoning training, by itself, is unlikely to have a large real-world impact. Current studies like FINGER and POINTER (Rosenberg et al., 2020) are emphasizing multi-component interventions, and another future direction would be to try to incorporate this kind of telehealth training into a broader intervention plan that not only trained other abilities, but also included ingredients like loneliness reduction, cardiovascular management, exercise, and stress reduction.

APPENDIX A
TELEHEALTH READINESS FORM AND TECHNICAL SUPPORT



Telehealth Readiness Form

Invitation status: | Survey options

Editing existing Record ID **BR Practice**

Event Name: **Pre-Test**

Record ID	BR Practice
What is your phone number? * must provide value	
What is your physical address?	Expand
Who is your emergency contact? * must provide value	
What is your emergency contact's phone number? * must provide value	
Do you have internet access? * must provide value	○ Yes ○ No reset
Is it faster than 10 Mbps download and 5 Mbps upload?	○ Yes ○ No reset

Do you have a computer with audio and visual capacity so we can hear and see each other? **CELL PHONES AND TABLETS ARE NOT ALLOWED.** * must provide value	○ Yes ○ No	reset
Do you have a computer mouse? * must provide value	○ Yes ○ No	reset
Would you be willing to be seen via video visits? * must provide value	○ Yes ○ No	reset
Do you have Zoom? * must provide value	○ Yes ○ No	reset
Have you used Zoom with family, friends, or healthcare providers? * must provide value	○ Yes ○ No	reset
Do you have a private and quiet space in your home? * must provide value	○ Yes ○ No	reset

To check your understanding of the Zoom features required for this course, we would like to ensure that you are able to:

(1) Share Screen
(2) Switch between Gallery and Speaker views
(3) Mute and Unmute
(4) Turn on and off video
(5) Use Chat
(6) Optimize Volume

Ability to Share Screen?

* must provide value

- Unable to Perform Action
- Perform with Assistance
- Perform without Assistance

reset

Ability to Switch Between Gallery and Speaker views?

* must provide value

- Unable to Perform Action
- Perform with Assistance
- Perform without Assistance

reset

Ability to Mute and Unmute?

* must provide value

- Unable to Perform Action
- Perform with Assistance
- Perform without Assistance

reset

Ability to Turn On and Off Video?

* must provide value

- Unable to Perform Action
- Perform with Assistance
- Perform without Assistance

reset

Ability to Use Chat?

* must provide value

- Unable to Perform Action
- Perform with Assistance
- Perform without Assistance

reset

Ability to Optimize Volume?

* must provide value

- Unable to Perform Action
- Perform with Assistance
- Perform without Assistance

reset

In addition to Zoom, you will be using Canvas to participate in the intervention. To be eligible for the intervention, you must be able to:

(1) Navigate to and from the Modules page
(2) Navigate between activities and modules

(3) Annotate Assignments

You are not eligible to participate if you cannot navigate Canvas. Have you completed the Canvas walkthrough?

- Yes
- No

reset

* must provide value

Ability to Annotate?

- Unable to Perform Action
- Perform with Assistance
- Perform without Assistance

reset

* must provide value

Form Status

Complete? Incomplete

APPENDIX B
ADAPTED TELEHEALTH-USABILITY QUESTIONNAIRE

Rating Scale: 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, and 5 = Strongly Agree

Usefulness

1. Telehealth improves my access to healthcare services?
2. Telehealth saves me time traveling to a hospital or specialist clinic?
3. These telehealth training sessions were beneficial?
4. These telehealth training sessions were appropriately challenging?

Ease of Use & Learnability

1. It was simple to use Zoom?
2. It was easy to learn to use Zoom?

Interface Quality

1. I like using the Zoom interface?
2. The Zoom interface is simple and easy to understand?
3. This system is able to do everything I would want it to be able to do?

Interaction Quality

1. I could easily talk to the trainer using the telehealth system?
2. I could hear the trainer clearly using the telehealth system?
3. I felt I was able to express myself effectively?
4. Using the telehealth system, I can see the trainer as well as if we met in person?

Reliability

1. I think the visits provided over the telehealth system are the same as in-person visits?
2. Whenever I made a mistake using the Zoom interface, I could recover easily and quickly?

Satisfaction and Future Use

1. I feel comfortable communicating with the trainer using the telehealth system?
2. Telehealth is an acceptable way to receive cognitive training?
3. I would use telehealth services again?
4. Overall, I am satisfied with this telehealth Zoom system?
5. These telehealth training sessions were enjoyable?

APPENDIX C
ACTIVE REASONING TRAINING AND TAYLOR TELEHEALTH SESSION-TO-SESSION DATA FORMS

ACTIVE Reasoning Training Session-To-Session Data Form

ACTIVE - REASONING TRAINING DATA FORM #904

A1. ID #: LABEL **A3.** Visit #: Z1 **A5.** Training Site Code: ______

A2. Form Version: 08/01/1998 **A4.** Replicate #: ______ **A6.** Training Site Name: ______

B.	**a.** SCHEDULED DATE	**b.** ACTUAL DATE	**c.** ACTUAL TIME	**d.** TRAINER ID #	**e.** EXPOSURE	**f.** ENVIRONMENT	**g.** RIGHTS	**h.** ATTEMPTS	**i.** HEARING DEVICE Y / N
1									1 / 2
2									1 / 2
3									1 / 2
4									1 / 2
5									1 / 2
6									1 / 2
7									1 / 2
8									1 / 2
9									1 / 2
10									1 / 2

e. EXPOSURE CODE:
- 12= Left 30 mins. early
- 13= Left 15 mins. early
- 14= Stayed 60 mins.
- 24= 15 mins. late
- 34= 30 mins. late

f. ENVIRONMENT CODE:

1	2	3	4	5	6	7
Very Difficult To Train			Adequate			Very Easy To Train

C. TRAINER PERCEPTION:

C1. Is your perception of training improvement consistent with practice exercise data?

YES……….1 NO……….2 →(C2)

C2. If NO, did subject…

do better……1 do worse….2

D. SUBGROUP: Basic…………………………….1
Standard…………………………..2

E. COMMENTS:

Taylor Telehealth Session-To-Session Data Form

Session To Session Form

Invitation status: Survey options

Editing existing Record ID **2**

Event Name: **Session-to-Session Data**

Record ID 2

Participant ID #:

Scheduled Date	Actual Date	Actual Time	Trainer ID #	Exposure	Environment	Rights	Attempts	Hearing Device Y / N
Today M-D-Y	Today M-D-Y	Now H:M						

Exposure Code	Condition
12 =	Left 30 mins. early
13 =	Left 15 mins. early
14 =	Stayed 60 mins.
24 =	15 mins. late
34 =	30 mins. late

Environment Code						
1	2	3	4	5	6	7
Very Difficult to Train			Adequate			Very Easy to Train

Is your perception of training improvement consistent with practice exercise data?

Comments:

Expand

APPENDIX D
ADVERTISEMENTS

Postcard

Flyer


Computer Video-conference Training in Problem Solving for Seniors
YOU MAY BE ELIGIBLE TO ENROLL IF:
• You are willing to complete videoconference problem solving training from home, for two days a week for 5 weeks
• You are 65+ and have never been diagnosed with Alzheimer's or dementia.
• You have a computer and internet.
Please email brad.taylor@ufl.edu or call the University of Florida (352) 273-5098 for more information.
Compensation will be provided.

APPENDIX E
SESSION SCRIPTS

Session 1 Script

1. Login early and help people login. I can demonstrate the process, on Firefox, by sharing my screen.

2. Introductions and get timer ready for all the tasks. Also introduce the RA! Tell RA that if anything happens to my internet, to just do their best to spur things along.

3. If folks are still not logged in, have them click the zoom link in their email and log in to canvas. I can demonstrate the process.

4. Tell them that we are using both Zoom and Canvas for this program, so what I recommend is, since for the most part you will want to see your screen to do the tasks on Canvas, I'll have you default to keeping Zoom minimized. When you click the minimize button on Zoom, it should make me a small box that you can actually move wherever you want by clicking and dragging me where you want me to go. However, some of you have a web version of Zoom, where Zoom is actually in a different browser tab, so in those cases, you'll just need to get comfortable switching from the Zoom tab on your Web Browser and the Canvas tab, also on your Web Browser. I mentioned I want you all to see your Canvas webpage for the most part, and I wanted to follow up and say that throughout the sessions and particularly these first few sessions, I'll be sharing my screen a lot to demonstrate how to use Canvas. You've all already used Canvas quite a bit during the pre-training assessment you may recall. And for those instances when I'm demonstrating things by sharing my screen, I'll have you make Zoom full screen so you can see exactly what I'm doing, then after I'm done sharing my screen, you can go right back to minimizing Zoom so you can see your Canvas page, without Zoom in the way. I want to also say that it's totally OK if we have some technical difficulties in this very first session; I have a feeling that as the sessions go on, this process will come more naturally, and I hope we all grow to feel more comfortable with this unique Zoom format!

5. Have them Click Session 1. Ask if they are ok with me recording just to send to my mentor for internal evaluation of how the first session is going.

6. Before we get started, I want to mention that the tasks you are being asked to perform will become progressively more complex, but at the same time you all will naturally get better at doing them because of all the practice you will receive. I also want to mention that this is meant to mimic a classroom-like setting, so in a

classroom at school, there is a limited amount of time, and the teacher is tasked with trying to get as many students through as much of the content as possible, but with time constraints. Likewise, we will try to adopt a mentality of let's just try to get as much done, but don't fret if you don't finish an activity in the time allowed. I have a stopwatch that I'll be using to keep us on a time schedule, but I don't want you to feel bad if you don't finish an entire activity. The point is to train your brain and just do your own personal best, not necessarily to finish every single activity, and we are also dealing with technological constraints, like internet speeds.

7. <u>4 minutes</u>: MENTAL EXERCISE. Share my screen and show them how to do this: Ok so now I'm going to have you all click where it just says "Mental Exercise". [read the first page]. Also, I want to mention that throughout our sessions, it can be expected that everyone will progress at different speeds. Therefore, each of you should only consider your own progress. We will not be doing a comparison between participants. Our interests are in how much each individual shows improvement. [have them click the next page button, read the second page]. So like it says, at the end of each session, we will have a little practice exercise, and that's just to see where you're at after each session we complete. It will be a pretty unfamiliar task, and it's meant to be a little puzzle to solve. I don't want you to worry about it, just be mentally prepared to do a little task at the end of today. [have them click the next page button, and read the third page]. [have them click the next page button]

8. <u>5 minutes</u>: PATTERNS IN EXERCISE ROUTINES. Read intro script for patterns in exercise routines. And say after the first paragraph "When we find a pattern in what we are doing, it helps us to lighten our memory burden. In the case of this exercise routine, instead of remembering 12 different steps, we only need to remember one **pattern** and then repeat it 3 times." Then read next two paragraphs. Demonstrate the pattern twice and say steps at the same time. Then have them do it with me WHILE saying the pattern. Read the last three paragraphs and have them 1) just do the exercise routine together, 2) think about the pattern, and 3) close your eyes and try to recall the 12 motions. [have them click the next page button]. This next page is what we call an annotated assignment, and you will recall doing this during the pre-training assessment, [share screen and make sure they all see my page and are all on this page, demonstrate entering full screen mode and using the toolbar at the top to zoom in and out]. I'll demonstrate how this will work, so recall that you need to click the paintbrush icon anytime we need to draw something, so I'm going to demonstrate by clicking the paintbrush icon and, as the instructions say, draw a slash after each pattern repetition. I'll go ahead and draw that now to demonstrate, and remember to click the check mark after every annotation you make to lock in or certify your answer. Then when you're done, click the exit full screen arrows that are like pointing at each other, scroll down and click submit assignment. Now I'm going to have you all try your hand at this. [make sure they're on the correct page and help them through this]. [click the next page button]

9. 5 minutes: OPTIONAL ACTIVITY: So you'll notice that some exercises are labeled as optional in Canvas, and that's to help with time management. If we are ahead of schedule, we can do some of these optional activities. For this one, I'm going to go ahead and have us skip it. [click the next page button, then click submit assignment on the next page and we will move on to the next page as well]. For optional breathe in breathe out activity, model the exercise routine twice by saying aloud each step, then have everyone join in, saying aloud each step, as I move my cursor over each step on the screen. Read final 3 paragraphs, doing what they say. [go on to next page for annotated assignment]. This next page is another annotated assignment, [share screen and make sure they all see my page and are all on this page, demonstrate entering full screen mode and using the toolbar at the top to zoom in and out]. I'll demonstrate how this will work, so recall that you need to click the paintbrush icon anytime we need to draw something, so I'm going to demonstrate by clicking the paintbrush icon and, as the instructions say, draw a slash after each pattern repetition. I'll go ahead and draw that now to demonstrate, and remember to click the check mark after every annotation you make to lock in or certify your answer [draw the slashes AND also underline the claps, if they ask why just say oh don't worry, we will go over that later, it just helps me see the pattern]. Then when you're done, click the exit full screen arrows that are like pointing at each other, scroll down and click submit assignment. Now I'm going to have you all try your hand at this. [make sure they're on the correct page and help them through this]. [click the next page button]

10. 6 minutes: FINDING THE PATTERN IN SCHEDULES. [Read first paragraph, ask if anyone has an example of a common memory problem of relevance, then read the other two paragraphs, and have them click the next page button]. [Make sure the annotated assignment is open for everyone but have them watch my screen share as I'll be demonstrating this whole annotated assignment before you do it, and demonstrate zooming in and out and clicking the paintbrush icon]. Then read instructions for example A or say There are five short schedules here. Four of them have the same kind of pattern. One is different. We can use these strategies to find the different one. Note that these are vertical, not horizonal. [demonstrate this while screensharing by scanning across the schedules and underlining any words in each schedule that are the same, while saying them out loud] Say "Not only can we see from the underlines which schedule is different, we can hear the difference when we say them out loud." [have them say them out loud]. "Which schedule is different?" Circle it on my screen. "We have used 3 **strategies** to find the pattern. 1) we **Scanned** the whole exercise, we **underlined** repeated words within a pattern, and we **said it aloud. Now,** I'm going to stop screen sharing and ask you to do Example A on your screen just as I demonstrated while we did it together. [when that is complete, have them click the next page button].

11. 4 minutes: Which schedule is different: instruct them to complete the 4 problems just as they did in the prior example, but I will demonstrate first [click screenshare] read instructions, do number 1 over screen share together as a group. Monitor until

finished. Let me know if you have any questions, don’t forget to click the check box after each of your answers to lock them in! [click the next page button]

12. <u>6 minutes</u>: Finding the pattern in longer schedules: share screen, read intro script paragraphs, direct their attention to example A, after saying underline, actually do it in example A, underlining, then say it aloud, reading the whole list and stressing the pattern and pausing slightly, then actually draw slashes, then say: “What's the pattern? You can **see** by the underlines that we have the first day repeated, then the next day once ... the following day repeated, the next day once ... the following day repeated, the next day once.” “When we said it out loud, we can **hear** the rhythm of the pattern. (Trainer, say it all again with voice emphasis). We put the **slash** where we hear the pattern begin to repeat itself.” [Then have them mark Example A on their own.] For example B share screen again, and say let’s use our strategies, scan it, then underline, and say aloud and slash for them, then stop screen share and have them do it. Same for c and d. [then have them click the next button]

13. <u>7 minutes</u>: Practice in finding patterns in schedules: no screen share, just make sure they’re on the page, and have them “do example A on their own screen and then go ahead and answer the next 5 questions on your own, we will go over them in a few minutes”. Monitor them. If many problems, model the steps for the first one, then have them do the next one. If problems still, just repeat helping them, and then having them try the next one. 4 minutes to work this one. 3 minutes to give pattern solutions. Do it myself and share answers “Tuesday Tuesday Wednesday slash Thursday Thursday Friday slash etc…

14. <u>4 minutes</u>: Finding the next item in the schedule: share screen and have them look at my screen, read intro script, Direct everyone’s attention to **Example A**. Let's find the pattern. Look over or scan the whole schedule. [I say the schedule aloud]. Remember one of our strategies is to underline any repeats . .. Hmmm . .. it looks like I need to underline everything . . . . [click the paintbrush icon and underline all the Mondays, Tuesdays, and Wednesdays] ...well actually that won't help me to see the pattern ...Sometimes this happens . .. when you say it aloud though, you hear the rhythm of the pattern . .. [say series aloud, then say “now that I say the pattern again, it is clear where the pattern repeats itself, and I can put slashes to denote where the pattern repeats itself]. [Put in the slashes where the pattern repeats]. “so now that we’ve figured out the pattern, what would come next in the series” [repeat saying aloud the pattern again] ... Wednesday is the next item if we continue the pattern. Write it on the line and show them how to click the free text annotation button “the T”, then have them try example b on their own.

15. <u>7 minutes</u>: Completing the schedule: read instructions and say “if you have difficulty doing one of the problems, just go on to the next one. We’ll go over them

all in a few minutes." Give them 5 mins to work, 2 mins to check. If problems, do first one over screen share.

16. 15 minutes: Group exercise: share my screen and tell them to just watch my screen, make sure they see me full screen. They don't need to download it. I'll download both documents and pull them up and screen share just the one document at first then both side by side when we get to that point. During the first example part that I walk them through, say "*The letters represent pills. On Sunday, this patient takes two of Pill A and one of Pill Bin the morning. On Sunday evening, he takes two of Pill A, and at bedtime on Sunday, he takes one Pill C Does everyone see that?" "What pills does he take on Thursday evening" Answer is Two A's. At bedtime? One C.we're going to make up a similar medication calendar for apatienet named mr. jones. Look at the prescription labels on these pages, when does Mr. Jones take medication A? Answer 3 times a day at meals. How many medications A's does he take at each meal? So where would you put one of mr. jones' a pills? Answer Sunday morning. Another? Sunday at noon. HAVE THEM EACH DOWNLOAD THE 2 DOCUMENTS AND WORK IN PAIRS, USING BREAKOUT ROOMS, ECNOURAGE ONE PERSON TO GIVE INSTRUCTIONS AND THE OTHER TO COMPLETE THE SCHEDULE. AFTER 10 MINUTES MAX, BRING THEM ALL BACK AND HAVE THEM HELP ME FILL OUT THE CHART WHILE SCREENSHARING WITH ALL OF THEM.*

8. You will need to fill out this schedule for Mr. Jones. Put an A, B. or C on the calendar when he should take the medication. Your calendar will look something like the sample one on page 16.

9. Encourage them to work in pairs—one giving instructions and one completing the schedule.

10. **Allow 10 minutes** for instruction and working.

11. **Allow 5 minutes** for participants to read answers. Have different participants read the schedule for each medication as trainer fills in the **blank schedule** on the flip chart.

17. 7 minutes: Practice Exercise on Letter Series: Say we are changing gears, and this is our final exercise of this session. Remember at the beginning of the session when I explained that we would need to do an assessment task, that's what we are going to do now. I'm going to ask you to do some problems that are different from the others we've worked on today. You will see, though, that the strategies we have been using will help here as well. I'll do example A for you on my screen. [do strategies of scanning, underlining, saying it out loud and have them say it out loud too, and put slashes in and then ask them what would come next and circle it for them]. THEN HAVE THEM DO EXAMPLE A ON THEIR SCREEN. Then say again I'll do example b for you on my screen as well. [do strategies of scanning, underlining, saying it out loud and have them say it out loud too, and put slashes in

and then ask them what would come next and circle it for them]. THEN HAVE THEM DO EXAMPLE B ON THEIR SCREEN.

18. 4.5 minutes twice: End of Session Reasoning Task. Read instructions. Remind them to use the strategies!

9. Participants may feel a great deal of anxiety at this point. It is important to continue the use of the physical exercise analogy. Explain that this timed exercise is not to "test" what they learned today. The purpose, as you explained at the beginning of the session, is to measure their strengths to see which types of material we should use in future sessions to meet their particular needs.

Compare this to asking them to do 20 sit-ups (or another appropriate physical exercise). They may not be able to do many—or any-- sit-ups right now, but after going through the whole exercise program, their ability to do sit-ups will improve because they have become stronger through practice, practice, practice. However, we need to know now how close to doing those "sit-ups" they are, so that we can develop the appropriate "exercise" program.

Assure them that whatever they can manage to do (or not do) is perfectly normal. Everyone will have problems with this because we haven't actually learned how to do them yet. Try to avoid describing the exercise as "hard". Use terms such as tricky, challenging, different, puzzling, or unusual. Encourage them to relax and have a little fun with it.

10. Give instructions for the timed exercise:

- They will be given 4-1/2 minutes to complete as many of the patterns as they can.

- Don't spend too much time on one problem-- go on and try another. They can come back and redo any that they skipped.

- Remind them that they should **use the strategies—scan, underline repeats, say aloud** (quietly so as not to disturb others), and **slash** between repetitions. They must find the pattern and then determine the letter which comes next in the series. It's ok to write the letter, but they must remember to **circle the one next letter in the series** in the answer column on the right.

11. **Time exercise for 4-1/2 minutes**. Trainer should not provide any assistance during this time. A whispered word of encouragement or clarification of instructions to an individual is appropriate.

Note: If an individual appears to be highly agitated and frustrated to the point that they are disturbing others (a rare occurrence), quietly tell them not to continue—you will take care of their assessment in another way.

12. End the exercise and collect all material for scoring later.

13. An **Optional Final Exercise** has been provided for use if trainer deems it necessary as a positive ending to the session. Allow no more than 3 minutes to complete.

19. Have them do more problems for 4.5 minutes but with red ink, instead of the default blue ink!

20. 3 minutes max: There is an optional thing at the end.

At the end of each session, remind them of when the next session is and ask if they can all make it and then spend time coordinating make-up sessions.

Mention that sometimes emails go to spam, is anyone not getting our emails?

Session 2 Script

1. Login early and help people login. I can demonstrate the process, on Firefox, by sharing my screen.

2. Introductions and get timer ready for all the tasks. Also introduce the RA! Tell RA that if anything happens to my internet, to just do their best to spur things along.

3. If folks are still not logged in, have them click the zoom link in their email and log in to canvas. I can demonstrate the process.

4. Explain Waiting Room and how I may have you wait until I'm ready to enter. I'll be there.

5. Have them Click Session 2. Possibly ask if they are ok with me recording just to send to my mentor for internal evaluation of how the first session is going.

6. <u>1 Minute</u>: TRAINING AND EVERYDAY ACTIVITIES. [Read the page. Then have them click the next button].

7. <u>2 Minutes</u>: TRAINING AND EVERYDAY ACTIVITIES: FINDING PATTERNS IN SCHEDULES PT. 2. [Read the page. Then have them click the next button].

8. <u>4 Minutes</u>: PATTERNS IN SHORT SCHEDULES (OPTIONAL). [This one is optional so feel out how it is going so far]. [share screen] "So this is actually an activity that we did during our last session. I'm going to read these instructions [read instructions]. Say we will use the strategies that we learned in Session 1, we start by scanning across all of the sets of schedules, and remember that these schedules are vertical, not horizontal [demonstrate scanning]. We then underline words in each of the schedules that repeat [underline repeats] and now we can already see which one is different, but let's go ahead and *say* all the schedules aloud so we can hear which one is different [have the group say each of the schedules]. So which one is different? [Read instructions and then have them do the 3 problems on their own, stop screensharing]. [click the next page button].

9. <u>8 minutes</u>: FINDING THE PATTERN IN LONGER SCHEDULES. [Share Screen]. Read the Intro script and the Example A script, then model the strategies: 1. Scan or look at every word in the schedule, 2. Underline repeated days or times, 3. Read aloud the schedule and ask "can you hear the pattern", and 4. Make slashes between repetitions of the pattern. Then ask "What's the pattern….[point to each segment as I describe the pattern: two of the same day, then one day/two of the same day, then one day/ two of the same day, then one day…"]. Have them do Example A on their own paper and say the schedule on the left has been done for

you, so just do the one the right. [Share my screen for **Example B**]. "I'll mark the pattern on my screen, then you'll do it on your screen". Go through steps of scanning, underlining, saying it aloud, and putting slashes where the pattern begins to repeat itself. "[point as I say] The pattern is two consecutive days, a double day, two consecutive days, a double day…etc.." [Then have them do it on their own]. [Share my screen for **Example C**]. "Example C is a little different. Let's do it together on the chart first." Demonstrate scanning, then underling and say "do you see already what's different about this one? We have days that repeat, but they are consecutive days instead of doubles. As we scanned down the list, we saw that Sunday and Monday kept repeating. So that's what we underline." "Let's say it aloud together" [say aloud together]. "The pattern is two consecutive days, then a single day; two consecutive days, then a single day; two consecutive days, a single day. I'll put slashes between each pattern repetition." [put in slashes]. [have them do example C on their own]. [Share my screen for **Example D**]. "Example D has another slightly different kind of pattern." [walk them thru scanning down the schedule, underlining repeats "hmm this is a little confusing, in this case, I want to say it aloud now so I can hear the repeats [say it aloud]. I can hear the rhythm, the days are consecutive and each day repeats, morning keeps repeating, afternoon keeps repeating…I *could* underling everything…but the most consistent repeat is morning, afternoon, so I'll just underline those words. The pattern is a day, morning, same day, afternoon / next day morning, next day afternoon / next day morning, next day afternoon. So I'm going to draw slashes to represent the pattern repetitions. [then have them do example D on their own and stop screen sharing]. [click the next page button].

10. <u>1 minute</u>: REMEMBER THE STRATEGIES-2. [Just read this to them]. [click the next page button].

11. <u>8 minutes</u>: PRACTICE IN FINDING PATTERNS IN SCHEDULES. [Screen Share and Read instructions and Model strategies for Example A, then have them work on the problems on the next pages. Encourage them to try as many as they can. 5 minutes of work time, 3 minutes to correct.

12. <u>4 minutes</u>: FINDING THE NEXT ITEM IN THE SCHEDULE. [Screen share and read the instructions and remind them that they will have to use both the paintbrush icon and the free text annotation "T" button for this. Do the strategies scan, underline, say aloud, and slashes, then say "the pattern is two consecutive days, a.m./two consecutive days, a.m. / two consecutive days, a.m./ [point as I say it], I have my two consecutive days, I need…. "a.m." write the answer on the blank and click enter. Then also demonstrate Example B the same way, scan, underline, say aloud, draw slashes, then ask "what comes next…the pattern is Tuesday, Wednesday, a single day/ Tuesday, Wednesday, a single day… I need to complete the last pattern, I have Tuesday, Wednesday… I need a single

day…which single day? What do the other single days have in common? Thursday, Friday, Saturday they're consecutive, so what's next…[write Sunday on the line].

13. 10 minutes: COMPLETING THE SCHEDULE. [Read instructions, encourage them to complete as many as they can.] Give them 8 minutes to work. 2 minutes to check.

14. 15 minutes: GROUP EXERCISE: RECYCLING SCHEDULE. [Note: Don't have them download it quite yet! Download both activities in advance, share my screen and read the instructions on the first page]. [demonstrate scanning and underlining key words AND draw slashes between Trash/Recycling/Brush/Furniture to show separation of the sections] "I'd like you to transfer the information on this page onto a calendar, much like you might do at home to help remember such a lot of information. So do you see how there are October, November, and December calendar pages…mark each day that trash will be picked up, each day that newspaper and cardboard will be picked up etc….so what day will you mark on your calendar that brush will be picked up? ANSWER: November 29th and December 27th. You can abbreviate any you wish, it's too small to write it all out. Have them download the two word documents and explain that they will type into the document. Explain that I will put them in pairs and move them to breakout rooms and that one person should read the schedule as the other records while sharing their screen to the other person.. Put them in pairs and into breakout rooms and have the RA move from breakout room to breakout room to help them with Word. ALLOW 7 MINUTES IN BREAKOUT ROOMS TO WORK, HAVE RA MOVE AROUND AND HELP THEM ALL. CAN QUICKLY GO OVER ANSWERS AT THE END.

15. 3 minutes: REPEATED PATTERNS IN DAILY ACTIVITIES (OPTIONAL). [Do only if time allowing]. [I can have people do this early if another group is still finishing the group/pair activity. If ahead of schedule, do as a group]. [Read intro script and insure they know that "00:00" represents any hour of the day. Have a volunteer screenshare and describe and mark the chart to reflect the 10-news-5-weather-10-sports-5-entertainment (or 10-5-10-5/news, weather, sports, entertainment) pattern. [have them then do the questions individually or as a group, reminding them of the "T" button]. [Check answers at the end].

16. 22 minutes: PRACTICE EXERCISE 2. Have them do example a and example b individually and make sure they understand how to do it. I can model strategies if necessary. [next page]. [read instructions and set timer for 4 minutes and have them do as many as they can. Say stop after 4 minutes, then give them 8 minutes to do it in green. After 8 minutes, have them switch to a red color, and we will circle

the correct answers to remaining problems, if not already correct with another color.

Preparation: Examples A and B on the chart.

Procedure:

1. Ask group to do **Example A**. Request answer. If necessary, demonstrate on chart. Use **Aux** A if necessary.
2. Have them do **Example B**. Use chart to model strategies if necessary. Use **Aux** B if necessary.
3. Set **timer for 4 minutes**. Read instructions at bottom of page 17.
4. At end of 4 minutes, have participants put pencils down.
5. Reset timer for **8 minutes**. Record data on score sheet.
6. Allow additional working time with **blue pencil** until the end of the 8 minutes.
7. During remaining time, all 20 problems should be corrected with **red pencil**. Trainer should give pattern solutions as well as the answer.

At the end of each session, remind them of when the next session is and ask if they can all make it and then spend time coordinating make-up sessions.

Note: For the first group, I ended after the first 4 minutes and didn't give more time for the extra 8 minutes of practice because they have to type the words in and no option to change colors, but I should've had them change colors and do another 8 more minutes.

Session 3 Script

1. Login early and help people login. I can demonstrate the process, on Firefox, by sharing my screen.

2. Introductions and get timer ready for all the tasks. Also introduce the RA! Tell RA that if anything happens to my internet, to just do their best to spur things along.

3. If folks are still not logged in, have them click the zoom link in their email and log in to canvas. I can demonstrate the process.

4. Have them Click Session 3. Ask if they are ok with me recording just to send to my mentor for internal evaluation of how the first session is going.

5. <u>2 minutes</u>: SCHEDULES FOR MONTHS, DAYS, AND TIMES: Read script. [Ask them about their own schedules and what patterns they may find in them. Their training session schedules have a good pattern to them].

6. <u>4 minutes</u>: FINDING THE NEXT ITEM IN SCHEDULES: [share screen and have them read Example A out loud, then use the strategies: scanning down the schedule, underlining repeats, saying it aloud and say, "the pattern is two consecutive days, then a single day", then put slashes in the pattern. Then ask them "what comes next…we have the two consecutive days, we need a single day. The other single days – Wednesday, Thursday, Friday- follow consecutive order, so we need the next consecutive day, which is… Saturday". Then have them do Example A on their own, and remind them to click the "T" button.] [For Example B, share screen and read the instructions, then stop screen sharing and have them do it on their own, then compare Example B answers after they're done.] [For Example C, read it aloud (scanning), then underline repeats, then say "it looks like this is one of those problems where everything needs underlined…" then say it aloud again and say "I'll just underline the second repeats, then put the slashes between pattern repetitions" "what comes next?" A: Sunday.] Then just have them click the next button, they don't need to do that one on their own.

7. <u>5 minutes</u>: EXERCISE: FINDING THE NEXT ITEM IN THE SCHEDULE. Have the group do the example problem on their own and stop there before doing the problems below. Have someone volunteer the answer. Then only demonstrate the pattern if people are confused "if necessary". Then have them do the problems below on their own. Allow 4 minutes of working time, including instructions, and then use the final 1 minute to check. Only read answer unless they need an explanation for any of them

8. 2 minutes: USING ABBREVIATIONS IN SCHEDULES. Just read it. Then have them read row 1 and 2 out loud, being sure to say the full days of the week for each abbreviation.

9. 7 minutes: FINDING THE PATTERN IN SCHEDULES WITH ABBREVIATIONS. [share screen and read script and stress that the exact same strategies are used with this horizontal question format]. Demonstrate scanning across the pattern, underlining repeats, then say aloud the patter (pointing to each as I say it), and then draw slashes where the pattern begins to repeat. Then ask "what comes next? What's the pattern? Monday twice, Tuesday once / Wednesday twice, Thursday once / Friday twice, Saturday once …. Sunday should be twice, so we need another Sunday. [optional: ask them "If we were to continue the pattern, what comes after Sunday?"]. Have them do Example A, B, C, and D on their own. Then check their annotations and answers. For Example D mention that it is one of those types that you could underline everything.

10. 15 Minutes: COMPLETING THE SCHEDULE. Give instructions and have them work the problems. Allow 9 minutes to work, then give them solutions to as many problems as time allows.

11. 12 minutes: GROUP EXERCISE: MEDICATION SCHEDULE. Read the script on the first page, and explain the medication calendar using the strategies like underlining column and row headers. Ask them when this person takes medication B, what about what do they take Friday evening? < just questions to make sure they comprehend it. Then say "now we are going to make up a similar medication calendar for a patient named Mrs. Baker. Look at these prescription labels. When does Mrs. Baker take Medication A? How many of A does she take?" "How many pills of medication B does the take? Two. Then on her calendar you would put two B's (BB) at those times when she should take them" "Now let's put Mrs. Baker's medications on the calendar's pages. Remember when we did this before? This is just the same, except that we will fill out the calendar for two weeks. [put people in breakout rooms and have me go into one breakout room, the RA into another and we share our screens, with the two activities side-by-side and have them help us fill it in. 12 minutes for instruction and break-out room work combined, 3 minutes for checking in on how it went.

12. Unsure: FINDING PATTERNS IN SCHEDULES (OPTIONAL). Read instructions and allow them to work the problems

13. Unsure: PRACTICE EXERCISE-3: EXAMPLES. [share screen, read instructions, demonstrate the strategies for Example A, then have them do Example A and Example B and then check answers AND strategy use.

14. <u>Unsure</u>: END OF SESSION TASK INSTRUCTIONS: Read the instructions and set timer for four minutes. Then have them click next.

15. <u>4 minutes + 4 minutes</u>: PRACTICE EXERCISE 3. 4 minutes in one color, then make sure they have “T” button selected, and then make sure they click outside of the box, then click a new color, and 4 more minutes for practice.

At the end of each session, remind them of when the next session is and ask if they can all make it and then spend time coordinating make-up sessions.

Session 4 Script

1. Login early and help people login. I can demonstrate the process, on Firefox, by sharing my screen.

2. Introductions and get timer ready for all the tasks. Also introduce the RA! Tell RA that if anything happens to my internet, to just do their best to spur things along.

3. If folks are still not logged in, have them click the zoom link in their email and log in to canvas. I can demonstrate the process.

4. Have them Click Session 4. Ask if they are ok with me recording just to send to my mentor for internal evaluation of how the first session is going.

5. <u>3 minutes</u>: LETTER PATTERNS. Read script. Mention examples of using alphabetical order.

6. <u>1 minute</u>: DOES THE NAME COME BEFORE OR AFTER?. Just read it.

7. <u>1 minute</u>: DOES THE NAME COME BEFORE OR AFTER? Pt. 2. Share screen and have them tell me what to circle.

8. <u>2 minutes</u>: DOES THE NAME COME BEFORE OR AFTER? Pt. 3. Just read instructions and have them work independently. Check answers after 2 minutes.

9. <u>4 minutes</u>: WHICH SET IS DIFFERENT? Share screen and Read intro paragraph. Demonstrate scanning and underlining for Example A and then ask “do you see which set is different…there is only one set that does not have a double letter” then say them aloud so we can hear the different pattern. [then have them do Example A on their own]. Demonstrate scanning and underlining for Example B. “in this case there aren’t any repeats within the sets, but something repeats across the sets. What is it? Xy. Let’s say all the sets aloud [say them]. Now which one is different? They all have an xy, but one set has an xy in the middle instead of at the end. So we circle it. [have them do Example B on their own].

10. <u>5 minutes</u>: PRACTICE ON LETTER SET PROBLEMS. Read instructions and allow them to work independently for 4 minutes. 1 minute to check answers.

11. <u>8 minutes</u>: FINDING PATTERN REPETITIONS IN LETTER SERIES. Read instructions and Share Screen and say “we are now going to use just letters instead of words or abbreviations. It may take a little time to get used to.” [demonstrate example A and say, “these are just like the letter sets you all just did; they are now just pushed together”, do scanning, underlining, saying aloud, then

say "the pattern is consecutive days repeated twice. Then the next two letters in the alphabet repeated [draw slashes]." "so what's the next letter? Circle that and then have them do Example A on their own. [demonstrate Example B and ask "what is the pattern" "yep, the patten always ends with x y. The first two letters of each repetition are in alphabetical order. We have ab cd ef gh, so what comes next? I. just I. It may be tempting to circle more than one but just circle the one next letter." [have them do Example B, C, and D on their own].

12. <u>1 minute</u>: REMEMBER THE STRATEGIES. Read page.

13. <u>15 minutes</u>: PRACTICE IN FINDING THE NEXT LETTER IN A SERIES. Read instructions and allow 10 minutes to work. Have them give answers after. I can go over the first one, saying what I underlined and where I slashed. Go over any of them that people struggled with.

14. <u>12 minutes</u>: GROUP EXERCISE: FOLLOWING DIRECTIONS IN DAILY ACTIVITIES. Download the three activities, screen share the first two side-by-side. Demonstrate scanning, underlining (the title, the numbers, and the pictures), and saying aloud, then put slashes between each step in the directions. Then lead them through Example A. Then put them in breakout rooms and encourage them to organize the questions by using the strategies. Have the RA's download activity 1 and 3 and split screen them and share them and remind them to have them read the question, then underline the key words in the question, scan the table and find the step number that answers the question and underline key words in the directions that answer the question. There is 1 chart per question, so you can switch to a fresh chart for each question.

15. <u>3 minutes</u>: PRACTICE EXERCISE ON LETTER SERIES: Read script before Example A. Have them do Example A on their own and Example B, then check answers.

16. <u>1 minute</u>: PRACTICE EXERCISE ON LETTER SERIES: INSTRUCTION PAGE. Read the instructions.

17. <u>4 + 7 minutes</u>: PRACTICE EXERCISE 4. Allow 4 minutes, then switch colors and give them 7 minutes with another color. Then check answers. Corrections made in another new color.

18. <u>Unsure</u>: EXTRA PRACTICE. Probably won't do (only do if not at 75 minutes yet). Only for those needing the challenge.

At the end of each session, remind them of when the next session is and ask if they can all make it and then spend time coordinating make-up sessions.

Session 5 Script

1. Login early and help people login. I can demonstrate the process, on Firefox, by sharing my screen.

2. Introductions and get timer ready for all the tasks. Also introduce the RA! Tell RA that if anything happens to my internet, to just do their best to spur things along.

3. If folks are still not logged in, have them click the zoom link in their email and log in to canvas. I can demonstrate the process.

4. Have them Click Session 5. Ask if they are ok with me recording just to send to my mentor for internal evaluation of how the first session is going.

5. <u>2 minutes</u>: DISCUSSION OF PATTERN SERIES. Read script.

6. <u>4 minutes:</u> REPEATED PATTERNS IN SCHEDULES. [share screen, read script and make sure they know that “00:00” represents any hour of the day. Point to each line as I talk. “the news always comes on at the top of the hour, it could be any hour – 12:00 or 1:00 or 8:00 – then at 10 minutes past the hour, the weather comes on. What time does the sports coverage begin? 15 past. Entertainment? 25 past. Then the pattern begins all over again at the half hour. How long is the news on? 10 minutes. How long is the weather on? 5 minutes. And you can also see that the pattern in the second 30 minutes is the same as in the first 30 minutes. “news 10, 5 weather, 10 sports, 5 entertainment”. Stop sharing screen and have them answer the questions on their own. Then check answers.

7. <u>6 minutes:</u> FINDING THE NEXT LETTER IN THE SERIES. Share screen and read the intro script. Demonstrate Example A. Have the group recite the strategies we’ve been using before they do the exercise. Demonstrate Example A and then ask them what’s the pattern here and what’s the next letter. Have them do Example B on their own, then check answers. Demonstrate Example C and say “sometimes it helps me to say the pattern out loud before I underline because I can hear myself saying some letters over so I’ll UNDERLINE them. Then I’ll put slashes between each repetition. I need to complete this last pattern repetition. The pattern is mno two letters / mno two different letters / etc. so what do I need to begin this pattern again. Model Example D and then ask “what’s the pattern here? Three letters, xo / three different letters, xo / etc… we need to begin the pattern again. we know we need a different letter. How do we know which one? Let’s see how the different letters relate to each other. [circle cde, fgh, ijk, and lmn] and say them out loud. “what’s next? O.”.

8. <u>10 minutes:</u> PRACTICE IN FINDING THE NEXT LETTER IN A SERIES. Give them 7 minutes to do the questions. Then go over answers for 3 minutes. Have them volunteer to share their answers.

9. <u>1 minute:</u> SKIP PATTERNS IN SERIES. Read the script.

10. <u>6 minutes</u>: SKIP PATTERNS IN A SERIES PART 2. [share screen and demonstrate Example A]. "let me show you a strategy for these skip patterns, when I'm doing a page of patterns I can tell it's a skip pattern because it's usually shorter and there don't seem to be any repeats when I scan across the series. When I say the series out loud [say it aloud], right away I notice that there's a day missing between each of these. What's missing? Tuesday. [place a tick mark by the Mo and then write Tu above the tick mark.] I call this a tick mark. What else? Thursday is missing [write tick mark and Th above and repeat across the series]. "I can also put slash marks between the pattern repetitions. [put slashes after ach tick mark]. "the pattern is "day, next day missing / day, next day missing / day, next day missing / ….so what goes in the blank? Sa." Have them do Example A and Example B on their own, then check answers. Share screen and demonstrate Example C. Read aloud then ask "is there anything missing here? Friday. [place tick mark and repeat for others in the series] then ask "what's the pattern? Two days, next day missing / two days, next day missing / 2 days, next day missing etc… add slashes. Repeat saying the series aloud and point to tick marks and say missing day it represents. "How do we begin the next pattern repetition? MO." Have them do Example C on their own.

11. <u>7 minutes:</u> FINDING PATTERNS IN SCHEDULES. Read instructions and allow 5 minutes to work, then check answers.

12. <u>5 minutes:</u> SKIPPED PATTERNS IN LETTER SERIES. [share screen and demonstrate Example A, and reinforce that they must skip a letter between I and K when deciding what comes next]. Have them do Example B then check answers. Then share screen and demonstrate Example C, then have them do Example C on their own.

13. <u>5 minutes:</u> PRACTICE ON LETTER SERIES WITH SKIPPED LETTERS. Give them 3-4 minutes to work then give solutions to all problems.

14. <u>15 minutes:</u> GROUP EXERCISE: READING AND UNDERSTANDING MEDICINE BOTTLE LABELS. Read introductory paragraph and indicate the areas on the chart as they are mentioned in the text…do Example A as a large group and also draw slashes between the three parts of the medication label, then do breakout rooms for the rest of the 4 problems (one medicine bottle label per question), with RA's screensharing.

15. <u>2 minutes:</u> PRACTICE EXERCISE: EXAMPLES. Have the group do Example A and then I do Example B while sharing my screen.

16. <u>1 minute</u>: PRACTICE EXERCISE INSTRUCTIONS. Read instructions.

17. <u>4 + 7 minutes</u>: PRACTICE EXERCISE: ANNOTATED PORTION. Give them 4 minutes to work then they change their paintbrush color and work for 7 additional minutes.

18. <u>No Time</u>: EXTRA PRACTICE. Only if people have more time and want more practice.

At the end of each session, remind them of when the next session is and ask if they can all make it and then spend time coordinating make-up sessions.

Session 6 Script

1. Login early and help people login. I can demonstrate the process, on Firefox, by sharing my screen.

2. Introductions and get timer ready for all the tasks. Also introduce the RA! Tell RA that if anything happens to my internet, to just do their best to spur things along.

3. If folks are still not logged in, have them click the zoom link in their email and log in to canvas. I can demonstrate the process.

4. Have them Click Session 6. Ask if they are ok with me recording just to send to my mentor for internal evaluation of how the first session is going.

5. <u>2 minutes</u>: CONGRATULATIONS. Read script.

6. <u>4 minutes</u>: FINDING THE NEXT ITEM IN LETTER SERIES. Read instructions, share screen and demonstrate Example A and say “I’ve scanned across the series, I don’t see any repeats…I suspect this might be a skip pattern. Let’s say it aloud. I’ll put tick marks in place of the missing letters [add tick marks as it is said aloud]”. “What’s the pattern? Letter, missing letter / letter, missing letter…so what comes next? W. Don’t forget to circle it on the right.” Have them do Example A on their own. Then share screen and demonstrate Example B and say “what should I underline?” [Do underlines]…then say “this is a good example of a problem where the say it aloud strategy is most helpful. [have them scan to find the answer and encourage them to say it aloud to hear what to underline.]”. “Let’s say it aloud, and you all tell me where to put the slashes…What comes next?” Then have them do Example B on their own. Then have them do Example’s C and D on their own and go over answers together.

7. <u>20 minutes</u>: PRACTICE ON LETTER SERIES. Read instructions. Have them work independently. 15 minutes to work, then give solutions to all problems for 5 minutes. Note to self: they may not need 15 minutes, if I had to guess.

8. <u>15 Minutes Combined</u>: GROUP EXERCISES: CHOOSING HEALTHY FOODS AND HOW TO FIND GOVERNMENT SERVICES. For Choosing Healthy Foods, share screen then I first read the instructions and use strategies to become familiar with the table, including scanning, underlining key words, and drawing slashes between the columns. Then demonstrate Example A. Then breakout rooms with RA’s leading one Room and me, another room, where we all make up meals. Bring everyone back to main room. Then for Government Services, I read the text and

do the strategies of scanning and underlining, be sure to underline bullet points I supposes. Then breakout rooms for questions 1-4.

9. <u>15 minutes</u>: COMPLETING THE SCHEDULE. Share screen and Demonstrate Number 1. And say “this is a good example of a pattern within a pattern. To determine what comes next, circle Tuesday, Wednesday, Thursday, and show that the pattern is completed with Friday. Then have them do the remaining problems individually. They work for 10 minutes, and we check answers for 5 minutes.

10. <u>2 minutes</u>: PRACTICE EXERCISE ON LETTER SERIES: EXAMPLES. Have them do Examples A and B and then go over answers.

11. <u>4 + 7 minutes</u>: PRACTICE EXERCISE: ANNOTATED PORTION. Read instructions, give them 4 minutes to work, then 7 minutes for additional practice in a different color.

12. Unsure: EXTRA PRACTICE-6. Have them work if they have extra time or want more practice.

At the end of each session, remind them of when the next session is and ask if they can all make it and then spend time coordinating make-up sessions.

Session 7 Script

1. Login early and help people login. I can demonstrate the process, on Firefox, by sharing my screen.

2. Introductions and get timer ready for all the tasks. Also introduce the RA! Tell RA that if anything happens to my internet, to just do their best to spur things along.

3. If folks are still not logged in, have them click the zoom link in their email and log in to canvas. I can demonstrate the process.

4. Have them Click Session 7. Ask if they are ok with me recording just to send to my mentor for internal evaluation of how the first session is going.

5. <u>2 minutes</u>: MENTAL TRAINING: read it.

6. <u>5 minutes</u>: FINDING THE NEXT ITEM IN LETTER SERIES. Share screen, read intro paragraph, demonstrate Example A and say "Example A is a skip pattern, so let's scan it, there are no repeats to underline, let's say it aloud, then put tick marks in place of missing letters, don't forget the tick mark after n. so what's the next letter? Yep, so you'd circle P and click the check mark". Then demonstrate Example B "this is a tricky one, let's do it together, scan the series, we're looking for things to underline, right away I'm having a little trouble with that, and since we do want to be able to do these fast, let's go right to the say aloud step [have group say aloud]. That's better. I could hear that pattern faster than I could see it, what's happening here? [put in slashes while saying the pattern]. ab / abc / abcd / abcd…e. As we progress across the series, each repetition repeats the alphabetical order of the previous repetition plus adds one more letter…an alphabetical progress. Go ahead and do yours" [have them do Example B and C "it looks really long, but if you follow the strategies, it won't be any different than the others" and D].

7. <u>13 + optional 4 minutes</u>: PRACTICE ON LETTER SERIES. Read instructions and have them work independently for 10 minutes, then check answers for 3 minutes. If anyone gets an answer wrong, just talk about it as a group. Optional but

*****6. **4-Minute Timed Speed Exercise.** If trainer determines that the group is ready to begin developing greater speed, the exercise can be done over again under a timed condition. Participants should be encouraged to work as quickly as they can—no data will be taken—and to try to combine or "skip" steps of the strategies when the pattern is readily apparent to them. They should also skip over problems when the pattern does not emerge almost immediately and then come back to those later. Care should be taken to make this a fun challenge that has no consequences. You won't even check—they can compare their answers to the one which was already done if they choose to. The emphasis is on speed.

This process will be repeated in later sessions, so it is an **option at this time**. It is recommended that this method be used when most of the group has begun to level off in the number completed in the final practice exercises and accuracy rates are high. The trainer will have also noticed that the participants have already begun to skip steps, such as underlining, and are finding the repetitive strategies a little tiresome. The trainer should continue to emphasize using the strategies, but beginning to combine the steps.

If the level of accuracy falls in today's final practice exercise, do not repeat this speed drill in Session 8. Try it again in Session 9.

8. <u>15 minutes</u>: GROUP EXERCISE: WHAT SHOULD I EAT? Share screen, Read intro script, and also correlate the lines in the chart to the slashes in the pattern problems. After demonstrating the strategies and before doing Example A as a group, also say "How much fat in a serving of franks?" [Draw lines on the chart from *fat* and *franks* and show how the point of intersection is where you find the answer]. Then do Example A together, then breakout rooms for the 6 questions.

9. <u>10 minutes</u>: COMPLETING THE SCHEDULE. Share screen, read instructions, and do number 1 together. Then have them work independently. 7 minutes to work, and 3 minutes to check answers and give solutions to all problems.

10. <u>Unsure, actually 5 minutes to complete and check answers</u>: FIND THE PATTERN (OPTIONAL). It also says they can work for 4 minutes then 2 minutes to check answers.

11. <u>2 minutes</u>: PRACTICE EXERCISE: EXAMPLES. Share screen and do Example A as a group. Then have them do Example B on their own (it says to do it as a group but I think on their own makes sense). Check answers.

12. <u>Unsure</u>: PRACTICE EXERCISE: INSTRUCTIONS-2. Just read it.

13. <u>4 + 7 minutes + 7 minutes if I want</u>: PRACTICE EXERCISE: ANNOTATED PORTION. 4 minutes in one color, 7 in another. 7 minutes to check answers (think of how that would work).

14. <u>Unsure</u>: EXTRA PRACTICE. If time allows.

At the end of each session, remind them of when the next session is and ask if they can all make it and then spend time coordinating make-up sessions.

Session 8 Script

1. Login early and help people login. I can demonstrate the process, on Firefox, by sharing my screen.

2. Introductions and get timer ready for all the tasks. Also introduce the RA! Tell RA that if anything happens to my internet, to just do their best to spur things along.

3. If folks are still not logged in, have them click the zoom link in their email and log in to canvas. I can demonstrate the process.

4. Have them Click Session 8. Ask if they are ok with me recording just to send to my mentor for internal evaluation of how the first session is going.

5. <u>2 minutes</u>: PRACTICE AND REVIEW: Read it.

6. <u>18 minutes</u>: PRACTICE ON LETTER SERIES. Share screen and Demonstrate the first problem. Have them work independently for 9 minutes, then 3 minutes to go over answers. If appropriate, re-do it and give them 4-minutes for speed (emphasizing combining strategies or doing them mentally, skipping the more difficult problems and returning to them last). Then if they want, we can go over answers again (optional).

7. <u>12 minutes</u>: GROUP EXERCISE: UNDERSTANDING A MEDICATION CHART. Read script, demonstrate strategies on chart, then do Example A. Then breakout rooms for questions 1-4.

8. <u>Unsure</u>: SHUTTLE EXPRESS RIDE FARES (OPTIONAL).

9. <u>Unsure</u>: LETTER SETS. Read it.

10. <u>Unsure</u>: WHAT IS THE ORDER NUMBER? (OPTIONAL).

11. <u>4 minutes</u>: WHICH SET IS DIFFERENT?. Share Screen and Demonstrate Example A and say “we will us our strategies but sometimes we don’t need to use them all or it may be helpful to use them in a different order. [scan] you might want to scan more than once…what do you see in almost all of the sets? VZ. Let’s underline all of those. So we want to know which set is different. Clearly, MCDB is different. We can see that it is different because it doesn’t have anything underlined. So we circle it.” Have them do Example A on their own. Demonstrate Example B and say “first we will scan it [scan] do you see anything that repeats? I don’t see a particular letter that repeats itself in most of the sets, but I do see that most of the sets have a double letter. I’ll underline the double letters [underline]. I

can see the one that's different. Say it aloud if you're not sure yet [say it aloud], we can hear that all but one set has a double letter. LL, CC, KK, ZZ, and we can see that all but one set have an underline." Have them do Example B on their own. Demonstrate Example C and say "it is a little tricky so scan it very carefully [scan it twice], do you see a pattern emerging…something that is the same about all but one of them? Maybe here we should say it aloud before we underline…does anyone see it? I've noticed that most of the four letter sets are in proper alphabetical order. [underline as I say…CDEF, MNOP, does ABDC fit? No. VWXY, GHIJ…which one is different…circle the answer]. Then have them do Example C on their own. Have them look at Example D on their screen. "First, we scan it [scan]; I'll give you a hint, it's a skip pattern. Does anyone we see the pattern yet?

12. <u>15 minutes</u>: PRACTICE ON LETTER SET PROBLEMS. Read instructions and have them work for 11 minutes, then check answers for 4 minutes. If people finish early, they can work on the next activity if I want.

13. <u>4 minutes</u>: WHAT COMES NEXT IN THE SCHEDULE (OPTIONAL). Read instructions and have them work for a few minutes or give one problem to each person and have them report back to the group.

14. <u>2 minutes</u>: PRACTICE EXERCISE: EXAMPLES. Have them do Example A and B on their own. Then check answers.

15. <u>1 minute</u>: PRACTICE EXERCISE 8: INSTRUCTIONS. Read them.

16. <u>4 + 10 +7 minutes</u>: PRACTICE EXERCISE. 4 minutes for scored portion. Then give them 10 minutes to work remaining problems. 7 minutes to give solutions to all problems.

17. <u>Unsure</u>. EXTRA PRACTICE. It is encouraged that everyone try these at this point in training.

At the end of each session, remind them of when the next session is and ask if they can all make it and then spend time coordinating make-up sessions.

Session 9 Script

1. Login early and help people login. I can demonstrate the process, on Firefox, by sharing my screen.

2. Introductions and get timer ready for all the tasks. Also introduce the RA! Tell RA that if anything happens to my internet, to just do their best to spur things along.

3. If folks are still not logged in, have them click the zoom link in their email and log in to canvas. I can demonstrate the process.

4. Have them Click Session 9. Ask if they are ok with me recording just to send to my mentor for internal evaluation of how the first session is going.

5. <u>1 minute</u>: PRACTICE AND REVIEW: Read it.

6. <u>12 minutes</u>: WHICH SET IS DIFFERENT? ANNOTATED PORTION. Share screen and Read instructions. Demonstrate Number 1 [scan…underline any repeated letters or patterns of letters…I don't see any…say them aloud…do you know the pattern? All but one has four letters in alphabetical order…ABDE is missing a letter.] Have the group work independently; **encourage speed**. Check answers as a group.

7. <u>15 + 10 minutes</u>: GROUP EXERCISE: HOW MUCH SHOULD I EAT and READING A BUS SCHEDULE. Share Screen and Demonstrate Strategies on the chart. Then do Example A as a large group. Then assign breakout rooms and do Questions 1-4 in breakout rooms. Then repeat for the second group exercise. (RYAN TAKES OVER AT SECOND GROUP EXERCISE)

8. <u>Unsure</u>: PRACTICE ON LETTER SERIES – ANNOTATED ASSIGNMENT. Read instructions and have them work. Then go over answers.

9. <u>3 Minutes</u>: REPEATED PATTERNS IN SCHEDULES (OPTIONAL) – ANNOTATED ASSIGNMENT: Either do it as a group or assign to individuals who complete the previous exercise early.

10. <u>2 Minutes:</u> PRACTICE EXERCISE: EXAMPLES A AND B: Read instructions. Have them do Examples A and B on their own. Then check answers.

11. <u>Unsure</u>: PRACTICE EXERCISE PART 2: INSTRUCTIONS: Just read them.

12. <u>4 + 9 minutes + 7 minutes</u>: PRACTICE EXERCISE – ANNOTATED ASSIGNMENT: 4 minutes for scored portion. Then give them 9 minutes to work

remaining problems in a different color. 7 minutes to give solutions to all problems, including answers and strategies. Make sure they don't change any of their answers.

13. Unsure: EXTRA PRACTICE – ANNOTATED ASSIGNMENT: It is recommended that everyone attempt some of the problems.

At the end of each session, remind them of when the next session is and ask if they can all make it and then spend time coordinating make-up sessions.

Tell them to bring their calendars to session 10 so we can schedule their post training assessments during session 10.

Session 10 Script

1. Login early and help people login. I can demonstrate the process, on Firefox, by sharing my screen.

2. Introductions and get timer ready for all the tasks. Also introduce the RA! Tell RA that if anything happens to my internet, to just do their best to spur things along.

3. If folks are still not logged in, have them click the zoom link in their email and log in to canvas. I can demonstrate the process.

4. Have them Click Session 10. Ask if they are ok with me recording just to send to my mentor for internal evaluation of how the first session is going.

5. <u>5 minutes</u>: OUR LAST SESSION AND PT.2 AND PT.3: Read it and have them flip through the pages.

6. <u>2 minutes</u>: WHICH SET IS DIFFERENT?: Read instructions and emphasize that the steps can be done mentally to increase speed. Have them do Examples A, B, and C and check answers; explain solutions if needed.

7. <u>15 minutes</u>: PRACTICE ON LETTER SET PROBLEMS: Read instructions and have them work, then check answers when it seems everyone is done or after 11 minutes, whichever comes first. Explain strategies for ones that caused difficulty. If appropriate, do the exercise again but only for 2 minutes timed.

8. <u>8 minutes</u>: PRACTICE IN FINDING PATTERNS IN SCHEDULES: Have the group do the example problem first and check answer to make sure no one is having difficulty. Then have everyone work independently. If a reminder of the process is needed, model strategies on the example. Check answers after ~6 or 7 minutes and explain strategies for any that caused difficulty.

9. <u>20 minutes</u>: GROUP EXERCISE: SENIOR ACTIVITIES: Share Screen and Read instructions and mark up the chart. Have them find a pattern for when bingo is played (either M W F 11 / OR 11 M W F OR MO WE FR 11-12 etc…) and repeat that process for each activity listed. Be lenient and accept any valid response. Breakout rooms and Screen Share for 4 questions.

Bingo	Mo We Fr 11-12
Crafts	Mo Fr 1-3
Strength Training	Mo We Fr 9-10
Bridge	Tu Th 1-3
Chair Exercises	Tu Th 9-10

10. <u>8 minutes</u>: SENIOR VAN SERVICE. Read script and have the group notice the patterns of the arrival times. Ask them like what time does the van arrive at the library and when does the last bus leave the senior housing complex to assess comprehension. Have them work independently and check answers after.

11. <u>5 minutes</u>: PRACTICE IN FINDING THE NEXT LETTER IN A SERIES: Read script an tell them this is a chance to warm up before the last timed practice on letter series. They should proceed at a relaxed pace, and we will go over the answers in a few minutes. Give 3-4 minutes to work. Then go over all answers, giving pattern solutions wherever needed.

12. <u>2 minutes</u>: PRACTICE EXERCISE: EXAMPLE PROBLEMS: Read instructions and have them do Examples A and B on their own. Go over answers.

13. <u>Unsure</u>: PRACTICE EXERCISE PT. 2: INSTRUCTIONS: Read them.

14. <u>4 + 8 + 6 minutes</u>: PRACTICE EXERCISE: ANNOTATED PORTION. 4 minutes for activity, then have them switch colors to answer remaining questions. Then 6 minutes to check answers.

15. <u>Unsure</u>: EXTRA PRACTICE – 10: Have them try these extra problems.

SEND EVERYONE THEIR QUESTIONNAIRES

SCHEDULE EVERYONE FOR THEIR POSTTRAINING ASSESSMENT AND STRESS THE IMPORTANCE OF ATTENDING IT. ALSO SEND EVERYONE THEIR FINAL QUESTIONNAIRES AND RECOMMEND TO THEM THAT THEY COMPLETE THEM PRIOR TO THE POSTTRAINING ASSESSMENT VISIT TO SAVE THEM SOME TIME.

Remind them that email reminders may come from an RA with a different zoom link than the one’s we use for our training sessions.

LIST OF REFERENCES

Andresen, E. M., Bowley, N., Rothenberg, B. M., Panzer, R., & Katz, P. (1996). Test-retest performance of a mailed version of the Medical Outcomes Study 36-Item Short-Form Health Survey among older adults. *Medical care*, 1165-1170.

Armfield, N. R., Bradford, M., & Bradford, N. K. (2015). The clinical use of Skype—For which patients, with which problems and in which settings? A snapshot review of the literature. *International Journal of Medical Informatics*, *84*(10), 737-742.

Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., . . . Tennstedt, S. L. (2002). Effects of cognitive training interventions with older adults: a randomized controlled trial. *Jama*, *288*(18), 2271-2281.

Ball, K., Edwards, J. D., Ross, L. A., & McGwin, J. G. (2010). Cognitive Training Decreases Motor Vehicle Collision Involvement of Older Drivers. *Journal of the American Geriatrics Society*, *58*(11), 2107-2113. https://doi.org/10.1111/j.1532-5415.2010.03138.x

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental psychology*, *23*(5), 611.

Baltes, P. B. (1997). On the incomplete architecture of human ontogeny: Selection, optimization, and compensation as foundation of developmental theory. *American psychologist*, *52*(4), 366.

Baltes, P. B., Kliegl, R., & Dittmann-Kohli, F. (1988). On the locus of training gains in research on the plasticity of fluid intelligence in old age. *Journal of Educational Psychology*, *80*(3), 392.

Baltes, P. B., & Lindenberger, U. (1988). On the range of cognitive plasticity in old age as a function of experience: 15 years of intervention research. *Behavior therapy*, *19*(3), 283-300.

Baltes, P. B., Sowarka, D., & Kliegl, R. (1989). Cognitive training research on fluid intelligence in old age: what can older adults achieve by themselves? *Psychology and aging*, *4*(2), 217.

Baltes, P. B., & Willis, S. L. (1982). Plasticity and enhancement of intellectual functioning in old age. In *Aging and cognitive processes* (pp. 353-389). Springer.

Billing, D. (2007). Teaching for transfer of core/key skills in higher education: Cognitive skills. *Higher education*, *53*(4), 483-516.

Blazer, D. G., Yaffe, K., & Karlawish, J. (2015). Cognitive aging: a report from the Institute of Medicine. *Jama*, *313*(21), 2121-2122.

Blieszner, R., Willis, S. L., & Baltes, P. B. (1981). Training research in aging on the fluid ability of inductive reasoning. *Journal of applied developmental psychology*, *2*(3), 247-265.

Blom, G. (1958). *Statistical estimates and transformed beta-variables* [Doctoral thesis, monograph, Almqvist & Wiksell]. DiVA. Stockholm. http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-75457

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

Boron, J. B., Turiano, N. A., Willis, S. L., & Schaie, K. W. (2007). Effects of cognitive training on change in accuracy in inductive reasoning ability. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *62*(3), P179-P186.

Bowers, J., Fredrickson, M., Hansen, B., & Errickson, J. (2022). RItools: Randomization inference tools. *R package version 0.1-18*.

Brailean, A., Huisman, M., Prince, M., Prina, A. M., Deeg, D. J. H., & Comijs, H. (2018). Cohort Differences in Cognitive Aging in the Longitudinal Aging Study Amsterdam. *The journals of gerontology. Series B, Psychological sciences and social sciences*, *73*(7), 1214-1223. https://doi.org/10.1093/geronb/gbw129

Burton, R. L., & O'Connell, M. E. (2018). Telehealth rehabilitation for cognitive impairment: randomized controlled feasibility trial. *JMIR research protocols*, *7*(2), e43.

Busija, L., Osborne, R. H., Nilsdotter, A., Buchbinder, R., & Roos, E. M. (2008). Magnitude and meaningfulness of change in SF-36 scores in four types of orthopedic surgery. *Health and quality of life outcomes*, *6*(1), 1-12.

Butler, M., McCreedy, E., Nelson, V. A., Desai, P., Ratner, E., Fink, H. A., . . . Brasure, M. (2018). Does cognitive training prevent cognitive decline? A systematic review. *Annals of internal medicine*, *168*(1), 63-68.

Bédard, M., Koivuranta, A., & Stuckey, A. (2004). Health impact on caregivers of providing informal care to a cognitively impaired older adult: rural versus urban settings. *Canadian Journal of Rural Medicine*, *9*(1), 15-23.

Chan, M. Y., Haber, S., Drew, L. M., & Park, D. C. (2016). Training older adults to use tablet computers: does it enhance cognitive function? *The Gerontologist*, *56*(3), 475-484.

Chandler, M. J., Locke, D. E., Crook, J. E., Fields, J. A., Ball, C. T., Phatak, V. S., . . . Smith, G. E. (2019). Comparative effectiveness of behavioral interventions on quality of life for older adults with mild cognitive impairment: a randomized clinical trial. *JAMA network open*, *2*(5), e193016-e193016.

Chandler, M. J., Parks, A., Marsiske, M., Rotblatt, L., & Smith, G. (2016). Everyday impact of cognitive interventions in mild cognitive impairment: a systematic review and meta-analysis. *Neuropsychology review*, *26*(3), 225-251.

Chiu, H.-L., Chu, H., Tsai, J.-C., Liu, D., Chen, Y.-R., Yang, H.-L., & Chou, K.-R. (2017). The effect of cognitive-based training for the healthy older people: A meta-analysis of randomized controlled trials. *PLoS One*, *12*(5), e0176742.

Cockrell, J. R., & Folstein, M. F. (2002). Mini-mental state examination. *Principles and practice of geriatric psychiatry*, 140-141.

Cook, S. E., Marsiske, M., Thomas, K. R., Unverzagt, F. W., Wadley, V. G., Langbaum, J. B., & Crowe, M. (2013). Identification of mild cognitive impairment in ACTIVE: algorithmic classification and stability. *Journal of the International Neuropsychological Society: JINS*, *19*(1), 73.

Craik, F. I., & Salthouse, T. A. (2011). *The handbook of aging and cognition*. Psychology press.

Cullum, C. M., Hynan, L., Grosch, M., Parikh, M., & Weiner, M. (2014). Teleneuropsychology: Evidence for video teleconference-based neuropsychological assessment. *Journal of the International Neuropsychological Society: JINS*, *20*(10), 1028.

Czaja, S., Moxley, J., & Rogers, W. A. (2021). Social support, isolation, loneliness, and health among older adults in the PRISM randomized controlled trial. *Frontiers in Psychology*, 4307.

Daihua, X. Y., Parmanto, B., Dicianno, B. E., & Pramana, G. (2015). Accessibility of mHealth self-care apps for individuals with spina bifida. *Perspectives in health information management*, *12*(Spring).

Dark-Freudeman, A., West, R. L., & Viverito, K. M. (2006). Future selves and aging: Older adults' memory fears. *Educational Gerontology*, *32*(2), 85-109.

Demiris, G., Doorenbos, A. Z., & Towle, C. (2009). Ethical considerations regarding the use of technology for older adults: The case of telehealth. *Research in gerontological nursing*, *2*(2), 128-136.

Denney, N. W. (1984). A model of cognitive development across the life span. *Developmental Review*, *4*(2), 171-191.

Diehl, M., Willis, S. L., & Schaie, K. W. (1995). Everyday problem solving in older adults: observational assessment and cognitive correlates. *Psychology and aging*, *10*(3), 478.

Dimidjian, S., Beck, A., Felder, J. N., Boggs, J. M., Gallop, R., & Segal, Z. V. (2014). Web-based mindfulness-based cognitive therapy for reducing residual depressive symptoms: an open trial and quasi-experimental comparison to propensity score matched controls. *Behaviour Research and Therapy*, *63*, 83-89.

Dixon, R. A., & Lachman, M. E. (2018). Risk and protective factors in cognitive aging: advances in assessment, prevention, and promotion of alternative pathways.

Dodge, H. H., Zhu, J., Lee, C.-W., Chang, C.-C. H., & Ganguli, M. (2014). Cohort effects in age-associated cognitive trajectories. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, *69*(6), 687-694.

Edwards, J. D., Xu, H., Clark, D. O., Guey, L. T., Ross, L. A., & Unverzagt, F. W. (2017). Speed of processing training results in lower risk of dementia. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, *3*(4), 603-611.

Ekstrom, R. B., Dermen, D., & Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests* (Vol. 102). Educational testing service Princeton, NJ.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods*, *41*(4), 1149-1160.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, *12*(3), 189-198.

Gandek, B., Sinclair, S. J., Kosinski, M., & Ware Jr, J. E. (2004). Psychometric evaluation of the SF-36® health survey in medicare managed care. *Health care financing review*, *25*(4), 5.

Gates, N. J., Rutjes, A. W., Di Nisio, M., Karim, S., Chong, L. Y., March, E., . . . Vernooij, R. W. (2019). Computerised cognitive training for maintaining cognitive function in cognitively healthy people in late life. *Cochrane database of systematic reviews*(3).

Gonda, J., & Abilities, S. K. S.-T. M. (1985). Test: Word Series Test. Palo Alto. In: CA: Consulting Psychologists Press.

Greenaway, M., Duncan, N., & Smith, G. (2013). The memory support system for mild cognitive impairment: randomized trial of a cognitive rehabilitation intervention. *International journal of geriatric psychiatry*, *28*(4), 402-409.

Greifer, N. (2022). *Matching Methods*. https://cran.r-project.org/web/packages/MatchIt/vignettes/matching-methods.html#optimal-pair-matching-method-optimal

Grosch, M. C., Weiner, M. F., Hynan, L. S., Shore, J., & Cullum, C. M. (2015). Video teleconference-based neurocognitive screening in geropsychiatry. *Psychiatry Research*, *225*(3), 734-735.

Gross, A. L., Rebok, G. W., Unverzagt, F. W., Willis, S. L., & Brandt, J. (2011). Cognitive predictors of everyday functioning in older adults: Results from the ACTIVE cognitive intervention trial. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *66*(5), 557-566.

Hansen, B. B., & Bowers, J. (2008). Covariate balance in simple, stratified and clustered comparative studies. *Statistical Science*, 219-236.

Hartman-Stein, P. E., & LaRue, A. Enhancing Cognitive Fitness in Adults A Guide to the Use and Development of Community-Based Programs.

Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberger, U. (2008). Enrichment effects on adult cognitive development: can the functional capacity of older adults be preserved and enhanced? *Psychological science in the public interest*, *9*(1), 1-65.

Hewitt, K. C., Rodgin, S., Loring, D. W., Pritchard, A. E., & Jacobson, L. A. (2020). Transitioning to telehealth neuropsychology service: Considerations across adult and pediatric care settings. *The Clinical Neuropsychologist*, *34*(7-8), 1335-1351.

Hill, N. T., Mowszowski, L., Naismith, S. L., Chadwick, V. L., Valenzuela, M., & Lampit, A. (2017). Computerized cognitive training in older adults with mild cognitive impairment or dementia: a systematic review and meta-analysis. *American Journal of Psychiatry*, *174*(4), 329-340.

Hill, P. L., Beck, E. D., & Jackson, J. J. (2021). Maintaining Sense of Purpose Following Health Adversity in Older Adulthood: A Propensity Score Matching Examination. *The Journals of Gerontology: Series B*.

Hintze, J. (2008). Power Analysis and Sample Size Software (PASS 2008). In. Kaysville, Utah, USA: www.ncss.com.

Holmqvist, M., Vincent, N., & Walsh, K. (2014). Web-vs telehealth-based delivery of cognitive behavioral therapy for insomnia: a randomized controlled trial. *Sleep medicine*, *15*(2), 187-195.

Horn, J. L. (1970). Organization of data on life-span development of human abilities. In *Life-span developmental psychology* (pp. 423-466). Elsevier.

Horn, J. L. (1978). Human ability systems. *Life-span development and behavior*.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, *26*, 107-129. https://doi.org/https://doi.org/10.1016/0001-6918(67)90011-X

Jenkinson, C., Lawrence, K., McWhinnie, D., & Gordon, J. (1995). Sensitivity to change of health status measures in a randomized controlled trial: comparison of the COOP charts and the SF-36. *Quality of Life Research*, *4*(1), 47-52.

Jobe, J. B., Smith, D. M., Ball, K., Tennstedt, S. L., Marsiske, M., Willis, S. L., . . . Leveck, M. D. (2001). ACTIVE: A cognitive intervention trial to promote independence in older adults. *Controlled clinical trials*, *22*(4), 453-479.

Kane, R. L., Butler, M., Fink, H. A., Brasure, M., Davila, H., Desai, P., . . . McCarten, J. R. (2017). Interventions to prevent age-related cognitive decline, mild cognitive impairment, and clinical Alzheimer's-type dementia.

Kelly, M. E., Loughrey, D., Lawlor, B. A., Robertson, I. H., Walsh, C., & Brennan, S. (2014). The impact of cognitive training and mental stimulation on cognitive and everyday functioning of healthy older adults: a systematic review and meta-analysis. *Ageing research reviews*, *15*, 28-43.

Khatri, N., Marziali, E., Tchernikov, I., & Shepherd, N. (2014). Comparing telehealth-based and clinic-based group cognitive behavioral therapy for adults with depression and anxiety: a pilot study. *Clinical Interventions in Aging*, *9*, 765.

Kinsella, K. G., & Phillips, D. R. (2005). Global aging: The challenge of success.

Klauer, K. J., & Phye, G. D. (2008). Inductive reasoning: A training approach. *Review of Educational Research*, *78*(1), 85-123.

Kueider, A. M., Parisi, J. M., Gross, A. L., & Rebok, G. W. (2012). Computerized cognitive training with older adults: a systematic review. *PloS one*, *7*(7), e40588.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*(4), 389-433.

Labouvie-Vief, G., & Gonda, J. N. (1976). Cognitive strategy training and intellectual performance in the elderly. *Journal of Gerontology*, *31*(3), 327-332.

Lachman, M. E., Baltes, P., Nesselroade, J. R., & Willis, S. L. (1982). Examination of personality-ability relationships in the elderly: The role of the contextual (interface) assessment mode. *Journal of Research in Personality*, *16*(4), 485-501.

Lachman, M. E., & Leff, R. (1989). Perceived control and intellectual functioning in the elderly: A 5-year longitudinal study. *Developmental Psychology*, *25*(5), 722.

Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social psychological and personality science*, *8*(4), 355-362.

Lampit, A., Hallock, H., & Valenzuela, M. (2014). Computerized cognitive training in cognitively healthy older adults: a systematic review and meta-analysis of effect modifiers. *PLoS medicine*, *11*(11), e1001756.

Lawson, D. W., Stolwyk, R. J., Ponsford, J. L., McKenzie, D. P., Downing, M. G., & Wong, D. (2020). Telehealth delivery of memory rehabilitation following stroke. *Journal of the International Neuropsychological Society*, *26*(1), 58-71.

Lezak, M. (1995). Concept formation and reasoning. *Neuropsychological assessment*, 602-649.

Lindenberger, U., & von Oertzen, T. (2006). Variability in cognitive aging: From taxonomy to theory. *Lifespan cognition: Mechanisms of change*, 297-314.

Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., . . . Cooper, C. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet*, *396*(10248), 413-446.

Lowenthal, M. F., & Berkman, P. L. (1967). *Aging and mental disorder in San Francisco: A social psychiatric study*. Jossey-Bass.

Margrett, J. A., & Willis, S. L. (2006). In-home cognitive training with older married couples: Individual versus collaborative learning. *Aging, Neuropsychology, and Cognition*, *13*(2), 173-195.

Marra, D. E., Hamlet, K. M., Bauer, R. M., & Bowers, D. (2020). Validity of teleneuropsychology for older adults in response to COVID-19: A systematic and critical review. *The Clinical Neuropsychologist*, *34*(7-8), 1411-1452.

McKeough, A., Lupart, J. L., & Marini, A. (2013). *Teaching for transfer: Fostering generalization in learning*. Routledge.

Monaghesh, E., & Hajizadeh, A. (2020). The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. *BMC Public Health*, *20*(1), 1-9.

Morris, J. C. (1996). Classification of dementia and Alzheimer's disease. *Acta Neurologica Scandinavica*, *94*(S165), 41-50.

Noack, H., Lövdén, M., Schmiedek, F., & Lindenberger, U. (2009). Cognitive plasticity in adulthood and old age: gauging the generality of cognitive intervention effects. *Restorative neurology and neuroscience*, *27*(5), 435-453.

Orata, P. T. (1928). The theory of identical elements, being a critique of Thorndike's theory of identical elements and a re-interpretation of the problem of transfer of training.

Organization, W. H. (2017). *Global diffusion of eHealth: making universal health coverage achievable: report of the third global survey on eHealth*. World Health Organization.

Organization, W. H. (2019). Risk reduction of cognitive decline and dementia: WHO guidelines.

Owsley, C., Sloane, M., McGwin Jr, G., & Ball, K. (2002). Timed instrumental activities of daily living tasks: relationship to cognitive function and everyday performance assessments in older adults. *Gerontology*, *48*(4), 254-265.

Parisi, J. M., Gross, A. L., Marsiske, M., Willis, S. L., & Rebok, G. W. (2017). Control beliefs and cognition over a 10-year period: Findings from the ACTIVE trial. *Psychology and Aging*, *32*(1), 69.

Parmanto, B., Lewis Jr, A. N., Graham, K. M., & Bertolet, M. H. (2016). Development of the telehealth usability questionnaire (TUQ). *International journal of telerehabilitation*, *8*(1), 3.

Patterson, C. (2018). World alzheimer report 2018.

Plemons, J. K., Willis, S. L., & Baltes, P. B. (1978). Modifiability of fluid intelligence in aging: A short-term longitudinal training approach. *Journal of Gerontology*, *33*(2), 224-231.

Poon, P., Hui, E., Dai, D., Kwok, T., & Woo, J. (2005). Cognitive intervention for community-dwelling older persons with memory problems: telemedicine versus face-to-face treatment. *International journal of geriatric psychiatry*, *20*(3), 285-286.

Raz, N. (2009). Decline and compensation in aging brain and cognition: Promises and constraints. *Neuropsychology review*, *19*(4), 411.

Rebok, G. W., Ball, K., Guey, L. T., Jones, R. N., Kim, H.-Y., King, J. W., . . . Unverzagt, F. W. (2014). Ten-Year Effects of the ACTIVE Cognitive Training Trial on Cognition and Everyday Functioning in Older Adults. *Journal of the American Geriatrics Society*, *62*(1), 16.

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. https://doi.org/10.1111/jgs.12607

Rosenberg, A., Mangialasche, F., Ngandu, T., Solomon, A., & Kivipelto, M. (2020). Multidomain interventions to prevent cognitive impairment, Alzheimer's disease, and dementia: From FINGER to World-Wide FINGERS. *The Journal of Prevention of Alzheimer's Disease*, *7*(1), 29-36.

Saczynski, J. S., Margrett, J. A., & Willis, S. L. (2004). OLDER ADULTS'STRATEGIC BEHAVIOR: EFFECTS OF INDIVIDUAL VERSUS COLLABORATIVE COGNITIVE TRAINING. *Educational Gerontology*, *30*(7), 587-610.

Saczynski, J. S., Willis, S. L., & Warner Schaie, K. (2002). Strategy use in reasoning training with older adults. *Aging, Neuropsychology, and Cognition*, *9*(1), 48-60.

Sala, G., & Gobet, F. (2019). Cognitive training does not enhance general cognition. *Trends in cognitive sciences*, *23*(1), 9-20.

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanism of a neglected phenomenon. *Educational psychologist*, *24*(2), 113-142.

Salthouse, T. A. (1993). Influence of working memory on adult age differences in matrix reasoning. *British Journal of Psychology*, *84*(2), 171-199.

Schaie, K. W. (1993). The Seattle Longitudinal Study: a thirty-five-year inquiry of adult intellectual development. *Zeitschrift fur Gerontologie*, *26*, 129-129.

Schaie, K. W. (1996). *Intellectual development in adulthood: The Seattle longitudinal study*. Cambridge University Press.

Schaie, K. W. (2005). What can we learn from longitudinal studies of adult development? *Research in human development*, *2*(3), 133-158.

Schaie, K. W., & Willis, S. L. (1986). Can decline in adult intellectual functioning be reversed? *Developmental Psychology*, *22*(2), 223.

Sherman, D. S., Mauser, J., Nuno, M., & Sherzai, D. (2017). The efficacy of cognitive intervention in mild cognitive impairment (MCI): a meta-analysis of outcomes on neuropsychological measures. *Neuropsychology review*, *27*(4), 440-484.

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. (2016). Do “brain-training” programs work? *Psychological Science in the Public Interest*, *17*(3), 103-186.

Sloane, P. D., Zimmerman, S., Suchindran, C., Reed, P., Wang, L., Boustani, M., & Sudha, S. (2002). The public health impact of Alzheimer's disease, 2000–2050: potential implication of treatment advances. *Annual review of public health*, *23*(1), 213-231.

Smith, G. E. (2016). Healthy cognitive aging and dementia prevention. *American Psychologist*, *71*(4), 268.

Smith, G. E., & Bondi, M. W. (2013). *Mild cognitive impairment and dementia: Definitions, diagnosis, and treatment*. Oxford University Press.

Spearman, C. (1923). *The nature of" intelligence" and the principles of cognition*. Macmillan.

Steinerman, J. R. (2010). Minding the aging brain: technology-enabled cognitive training for healthy elders. *Current Neurology and neuroscience reports*, *10*(5), 374-380.

Stine-Morrow, E. A. (2007). The Dumbledore hypothesis of cognitive aging. *Current Directions in Psychological Science*, *16*(6), 295-299.

Stine-Morrow, E. A., & Basak, C. (2011). Cognitive interventions. In *Handbook of the psychology of aging* (pp. 153-171). Elsevier.

Stuart, E. A., King, G., Imai, K., & Ho, D. (2011). MatchIt: nonparametric preprocessing for parametric causal inference. *Journal of statistical software*.

Taylor, B., Kliegel, M., Willis, S., Rebok, G., & Marsiske, M. (2020). Predicting reasoning change trajectories in ACTIVE: A mixed modeling approach. Published Abstract. In (Vol. 26, pp. I-196). Final Program 2020 of the International Neuropsychological Society: Journal of the International Neuropsychological Society.

Taylor, B. P., Rebok, G. W., & Marsiske, M. (2022). Good clinical practice improves rigor and transparency: Lessons from the ACTIVE trial. *Psychology and aging*, *37*(1), 43.

Thomas, K. R., Rebok, G. W., & Willis, S. L. (2019). Advanced Cognitive Training for Independent and Vital Elderly (ACTIVE). In D. Gu & M. E. Dupre (Eds.), *Encyclopedia of Gerontology and Population Aging* (pp. 1-7). Springer International Publishing. https://doi.org/10.1007/978-3-319-69892-2_1075-1

Thurstone, L. (1938). The perceptual factor. *Psychometrika*, *3*(1), 1-17.

Thurstone, L. L., & Thurstone, T. G. (1947). SRA primary mental abilities.

Tombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: a comprehensive review. *Journal of the American Geriatrics Society*, *40*(9), 922-935.

Tuckson, R. V., Edmunds, M., & Hodgkins, M. L. (2017). Telehealth. *New England Journal of Medicine*, *377*(16), 1585-1592.

Unverzagt, F. W., Guey, L. T., Jones, R. N., Marsiske, M., King, J. W., Wadley, V. G., . . . Tennstedt, S. L. (2012). ACTIVE cognitive training and rates of incident dementia. *Journal of the International Neuropsychological Society: JINS*, *18*(4), 669.

Wadley, V. G., Crowe, M., Marsiske, M., Cook, S. E., Unverzagt, F. W., Rosenberg, A. L., & Rexroth, D. (2007). Changes in everyday function in individuals with psychometrically defined mild cognitive impairment in the Advanced Cognitive Training for Independent and Vital Elderly Study. *Journal of the American Geriatrics Society*, *55*(8), 1192-1198.

Ware, J. E., Kosinski, M., & Keller, S. (1994). *SF-36 physical and mental health summary scales: a user's manual*. Health Assessment Lab.

Ware Jr, J. E., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical care*, 473-483.

Willis, S., & Schaie, K. (1994). Cognitive training in the normal elderly. *Plasticité cérébrale et stimulation cognitive*, 91-113.

Willis, S. L., Blieszner, R., & Baltes, P. B. (1981). Intellectual training research in aging: Modification of performance on the fluid ability of figural relations. *Journal of Educational Psychology*, *73*(1), 41.

Willis, S. L., & Caskie, G. I. (2013). Reasoning training in the ACTIVE study: how much is needed and who benefits? *Journal of aging and health*, *25*(8_suppl), 43S-64S.

Willis, S. L., Jay, G. M., Diehl, M., & Marsiske, M. (1992). Longitudinal change and prediction of everyday task competence in the elderly. *Research on aging*, *14*(1), 68-91.

Willis, S. L., & Schaie, K. W. (1986). Training the elderly on the ability factors of spatial orientation and inductive reasoning. *Psychology and aging*, *1*(3), 239.

Willis, S. L., & Schaie, K. W. (2009). Cognitive training and plasticity: theoretical perspective and methodological consequences. *Restorative neurology and neuroscience*, *27*(5), 375-389.

Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., . . . Stoddard, A. M. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *Jama*, *296*(23), 2805-2814.

Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., . . . Active Study Group, f. t. (2006). Long-term Effects of Cognitive Training on Everyday Functional Outcomes in Older Adults. *JAMA*, *296*(23), 2805-2814. https://doi.org/10.1001/jama.296.23.2805

Winocur, G., Craik, F. I., Levine, B., Robertson, I. H., Binns, M. A., Alexander, M., . . . Mchugh, T. (2007). Cognitive rehabilitation in the elderly: Overview and future directions. *Journal of the International Neuropsychological Society: JINS*, *13*(1), 166.

Wolinsky, F. D., Mahncke, H. W., Weg, M. W. V., Martin, R., Unverzagt, F. W., Ball, K. K., . . . Tennstedt, S. L. (2009). The ACTIVE Cognitive Training Interventions and the Onset of and Recovery from Suspected Clinical Depression. *The Journals of Gerontology: Series B*, *64B*(5), 577-585. https://doi.org/10.1093/geronb/gbp061

Wolinsky, F. D., Unverzagt, F. W., Smith, D. M., Jones, R., Wright, E., & Tennstedt, S. L. (2006). The Effects of the ACTIVE Cognitive Training Trial on Clinically Relevant Declines in Health-Related Quality of Life. *The Journals of Gerontology: Series B*, *61*(5), S281-S287. https://doi.org/10.1093/geronb/61.5.S281

Wolinsky, F. D., Vander Weg, M. W., Martin, R., Unverzagt, F. W., Willis, S. L., Marsiske, M., . . . Tennstedt, S. L. (2010). Does cognitive training improve internal locus of control among older adults? *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *65*(5), 591-598.

Yam, A., Gross, A. L., Prindle, J. J., & Marsiske, M. (2014). Ten-year longitudinal trajectories of older adults' basic and everyday cognitive abilities. *Neuropsychology*, *28*(6), 819.

Yu, D. X., Parmanto, B., Dicianno, B. E., Watzlaf, V. J., & Seelman, K. D. (2017). Accessibility needs and challenges of a mHealth system for patients with dexterity impairments. *Disability and Rehabilitation: Assistive Technology*, *12*(1), 56-64.

Zahodne, L. B., Meyer, O. L., Choi, E., Thomas, M. L., Willis, S. L., Marsiske, M., . . . Parisi, J. M. (2015). External locus of control contributes to racial disparities in memory and reasoning training gains in ACTIVE. *Psychology and Aging*, *30*(3), 561.

Zelinski, E. M., Dalton, S. E., & Smith, G. E. (2011). Consumer-based brain fitness programs. In *Enhancing Cognitive Fitness in Adults* (pp. 45-66). Springer.

BIOGRAPHICAL SKETCH

Brad Taylor graduated from Hanover College, in Hanover, IN with a bachelor's degree in neuroscience in 2015. Immediately following, Brad enrolled in a master's degree in psychology at San Diego State University, under the mentorship of Dr. Paul Gilbert, studying the verbal memory differences between individuals with Parkinson's disease and Huntington's disease. Mr. Taylor matriculated into the doctoral program in clinical and health psychology at the University of Florida in 2017. He is currently pursuing his doctorate in clinical psychology, with a specialization in neuropsychology. His research interests involve open science best practices and interventions focused on delaying cognitive decline in older adults. He will receive his degree in the summer of 2023 after completing his predoctoral internship at the Connecticut VA – West Haven.