

Journal of Aging and Health

<http://jah.sagepub.com/>

Reasoning Training in the ACTIVE Study: How Much Is Needed and Who Benefits?

Sherry L. Willis and Grace I. L. Caskie
J Aging Health 2013 25: 43S
DOI: 10.1177/0898264313503987

The online version of this article can be found at:
http://jah.sagepub.com/content/25/8_suppl/43S

Published by:



<http://www.sagepublications.com>

Additional services and information for *Journal of Aging and Health* can be found at:

Email Alerts: <http://jah.sagepub.com/cgi/alerts>

Subscriptions: <http://jah.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

>> [Version of Record](#) - Jan 2, 2014

[What is This?](#)

Reasoning Training in the ACTIVE Study: How Much Is Needed and Who Benefits?

Journal of Aging and Health
25(8S) 43S–64S

© The Author(s) 2013

Reprints and permissions:
sagepub.com/journalsPermissions.nav

DOI: 10.1177/0898264313503987

jah.sagepub.com



Sherry L. Willis, PhD¹
and Grace I. L. Caskie, PhD²

Abstract

Objective: To characterize change through 5-year follow-up, associated with training, booster, adherence, and other characteristics. **Methods:** Sample included all individuals randomly assigned to reasoning training (N = 699). Piecewise latent growth modeling was used to examine trajectory of performance on outcome measures. **Results:** Training resulted in improved reasoning performance through Year 5. A significant third annual booster effect was one-half the size of the training effect. Training adherence resulted in greater training effects. Higher education, Mini-Mental State Exam (MMSE), better health, and younger age related to higher baseline performance. Higher MMSE was related to larger training effects, larger linear slopes, and smaller booster effects. Significant functional outcomes included a training effect for complex reaction time (CRT), and first annual booster effects for the CRT and observed tasks of daily living. **Discussion:** Initial training gain was comparable with magnitude of age-related cognitive decline over 5 years with no training. Neither age nor gender predicted training or booster effects, indicating the generality of training effects across age (65-90 years).

¹University of Washington, Seattle, WA, USA

²Lehigh University, Bethlehem, PA, USA

Corresponding Author:

Sherry L. Willis, PhD, Department of Psychiatry and Behavioral Sciences, University of Washington Medical School, 180 Nickerson St, Suite 260, Seattle, WA 98109, USA.

Email: oldage@u.washington.edu

Keywords

ACTIVE trial, aging, cognitive training, reasoning training, training adherence, transfer of training, intervention studies

The primary objective of the ACTIVE (Advanced Cognitive Training for Independent and Vital Elderly) randomized controlled trial was to test the effectiveness and durability of three distinct cognitive interventions in improving the performance of elderly persons on basic measures of cognition and on measures of cognitively demanding tasks of daily living (Jobe et al., 2001; S. L. Willis et al., 2006). Prior ACTIVE analyses have focused on comparison of training effects to performance of a no-contact control group (Ball et al., 2002; S. L. Willis et al., 2006). This article focuses on the reasoning training group only, examining the performance trajectory over 5 years and the effects of booster, training adherence, and covariates in greater detail.

Participants in ACTIVE were randomly assigned to one of four groups: training on verbal episodic memory, reasoning ability, speed of processing, or a no-contact control group. Participants were reassessed immediately after training and at first, second, third, and fifth annual follow-up (Jobe et al., 2001). Prior to first and third annual follow-up, a random subsample of each training group received booster training on the target ability. Significant training effects in comparison with the no-contact control have been found immediately after training and at all follow-up occasions, through the fifth annual follow-up (S. L. Willis et al., 2006). Immediately after training, 74% of reasoning, 86% of speed of processing, and 26% of memory participants showed reliable cognitive improvement (Ball et al., 2002). Booster training for reasoning and speed of processing training groups produced significantly better performance (net of initial training effect) on their targeted cognitive abilities (S. L. Willis et al., 2006).

The major aim of ACTIVE was to examine whether cognitive training was associated with maintenance of functioning on tasks of daily living. Both self-reported and performance-based measures of daily function were included in ACTIVE (Diehl et al., 2005). At the fifth annual follow-up, the reasoning training group reported significantly less difficulty in the instrumental activities of daily living than the control group (S. L. Willis et al., 2006). The booster training for the speed of processing group showed a significant effect on the performance-based functional measure of everyday speed of processing.

Reasoning, as well as memory and speed of processing, was selected for the ACTIVE trial because prior research has shown them to be important basic abilities in cognitive functioning (K. Schaie, 2005), and because of previous successful interventions on these abilities in smaller scale training studies (Ball, Edwards, & Ross, 2007; Kramer & Willis, 2002; Smith et al., 2009; Verhaeghen, Marcoen, & Goossens, 1992). Inductive reasoning ability has been identified as a component of executive functioning (Lezak, 1995) and has been shown to be closely related to working memory ability (Salthouse, Fristoe, McGuthry, & Hambrick, 1998). Reasoning is considered one of the primary markers of fluid intelligence and has been found in longitudinal research to exhibit relatively early age-related decline, beginning in the mid-60s (K Schaie, 2005). Prior research has shown reasoning ability to account for significant individual variability in self-reported everyday functioning and in performance-based assessment of instrumental activities of daily living (Allaire & Marsiske, 2002; Diehl, Willis, & Schaie, 1995; M. Marsiske & Willis, 1995).

Earlier, smaller scale training studies involving no contact controls had found training effects on reasoning ability similar to the findings of the ACTIVE trial. The early ADEPT studies (Blieszner, Willis, & Baltes, 1981) reported immediate training effects when rural elderly were trained on strategies associated with reasoning ability; follow-up of the ADEPT sample showed maintenance of effects 7 years after training (S. Willis & Nesselroade, 1990). The training studies within the Seattle Longitudinal Study (SLS; Schaie & Willis, 1986; S. Willis & Schaie, 1994) indicated that both individuals showing prior longitudinal decline on reasoning and individuals with no prior decline could improve as a function of reasoning training; these training effects were replicated over three waves of training with successive adult birth cohorts (S. Willis & Schaie, 1994). Approximately one-half of SLS reasoning training participants showed significant pre- and posttest improvement; after training, 40% of those showing prior reasoning decline were performing at or above their level of performance 14 years prior to training. Training effects have also been found at the level of individual reasoning measures and at the level of a reasoning ability factor, defined by confirmatory factor analyses (S. Willis & Schaie, 1994).

Reasoning training has focused on strategies associated with problem solving involving patterns or rules. Reasoning training participants had higher usage of strategies than a comparison group and higher strategy use was associated with reasoning training gain (Saczynski, Willis, & Schaie, 2002). Collaborative training of couples on reasoning ability has been shown

to be effective; older adults using a self-directed training protocol can improve their reasoning ability without the presence of a trainer (Margrett & Willis, 2006; Saczynski, Margrett, & Willis, 2004). The magnitude of training effects has been found to vary with cognitive risk and dementia status (Boron, Willis, & Schaie, 2007). Although these prior training studies had demonstrated the efficacy of reasoning training, the findings of these earlier studies were limited by the lack of intent-to-treat analyses, by small and homogeneous samples, and in some cases incomplete random assignment (Stine-Morrow & Basak, 2011).

Given the ACTIVE findings of 5-year durability of training effects and some transfer to everyday functioning, there has been considerable interest in further examination of the characteristics of individuals profiting from reasoning training and of issues of dosing, including adherence with training and added effects of booster training. Prior reports of findings of training effects from the ACTIVE trial (Ball et al., 2002; S. L. Willis et al., 2006) have compared each training intervention with the control group, using a repeated-measures, mixed-effects model. In this article, we use piecewise growth models from baseline to the fifth annual follow-up to examine the 5-year trajectory separately for the reasoning training group. Although only the reasoning composite score was used in the prior articles to represent the proximal outcome of the reasoning training, this article reports findings for the composite and three individual reasoning tests (letter series, letter sets, word series; see McArdle & Prindle, 2008). Three major questions were addressed: (1) What are the trajectories for the reasoning training group through the fifth annual follow-up on the cognitive and functional outcomes? (2) Does intervention dosing (i.e., adherence with training regimen and boosters) influence outcomes? and (3) Are covariates significant predictors of training effects?

Method

Participants

The sample used for this study included all individuals randomly assigned to reasoning training ($N = 699$; $M = 162$, $F = 537$) and included an analyses consistent with the intention-to-treat principle. At baseline, this sample ranged in age from 65 to 91 years ($M = 73.53$, $SD = 5.76$), with an average education of 13.50 years ($SD = 2.69$, range = 4-20 years), and mean baseline Mini-Mental State Exam (MMSE) of 27.26 (range = 23-30; $SD = 2.02$). Demographic characteristics for those in the booster and no booster groups are presented in Table 1.

Table 1. Demographic Characteristics of Reasoning Training Participants ($N = 699$).

| | Booster ($n = 370$) | No booster ($n = 329$) |
|--------------------|-----------------------|--------------------------|
| Age | | |
| <i>M</i> | 73.11 | 74.02 |
| Range | 65-88 | 65-91 |
| Years of education | | |
| <i>M</i> | 13.46 | 13.55 |
| Range | 6-20 | 4-20 |
| MMSE | | |
| <i>M</i> | 27.38 | 27.12 |
| Range | 23-30 | 23-30 |
| Female | | |
| <i>n</i> | 283 | 254 |
| % | 76.5 | 77.2 |
| Good health | | |
| <i>n</i> | 303 | 255 |
| % | 81.9 | 77.5 |

Note. MMSE = Mini-Mental State Exam.

The attrition information for the ACTIVE sample is described in S. L. Willis et al. (2006). Based on the intent-to-treat design of a randomized clinical trial, 30 subjects (4.3%) randomized to reasoning training but who received no reasoning training are included in the 699 participants analyzed for this study. Of the 699 persons included in this study, 625 (89.4%) completed 8 out of 10 initial training sessions and were considered compliant with training. Approximately 60% of participants ($n = 370$) who were compliant with initial reasoning training were randomly assigned to reasoning booster training at Year 01. Of individuals participating in first annual booster training ($n = 305$), 98.6% ($n = 301$) completed three out of four booster sessions (i.e., were compliant); 66 persons assigned to booster training did not participate in Year 01 booster. Of individuals participating in Year 03 booster ($n = 250$), 97.2% ($n = 243$) were compliant; 121 originally assigned to booster training did not receive third annual booster. At Year 05, 469 (67%) reasoning participants were assessed; 64 (9%) had died; 126 (18%) had withdrawn due to health, moving away, caregiving, and so on; 40 (5.7%) had been administratively withdrawn due to loss to follow-up. The 5-year retention rate for reason trained participants was comparable with the 67% retention rate for the total ACTIVE sample at 5 years.

Study Design and Protocol

The ACTIVE trial was a multisite, single-blind randomized controlled clinical trial. The ACTIVE sample consisted of older adults aged 65 to 94 years living independently who were recruited from senior housing, community centers, hospitals, and clinics in Birmingham, AL, Detroit, MI, Boston, MA, Indianapolis, IN, central PA, and Baltimore, MD (Ball et al., 2002; Jobe et al., 2001). Participants were excluded from the trial if there was evidence at the time of enrollment of substantial cognitive (MMSE score < 23 or a diagnosis of dementia), visual (self-report of excessive difficulty reading newspaper print or visual test resulting in a score worse than 20/50), or functional decline (difficulty with activities of daily living [ADLs]), or a diagnosis of certain medical conditions that would predispose one to possible functional decline or mortality, such as short-lived cancers or stroke. Study procedures were approved by the institutional review boards at the collaborating institutions and all subjects gave informed consent to participate.

Procedure

Potential participants were first telephone-screened to determine eligibility. Eligible participants consented to three in-person baseline assessments (two individual and one group session) of cognitive, sensory, and functional abilities. Participants were also interviewed regarding their health, medication use, and everyday activities. After completing all eligibility and baseline assessments, participants were randomized to either a no-contact control group ($n = 704$), speed of processing training group ($n = 712$), memory training group ($n = 712$), or reasoning training group ($n = 705$). However, due to a protocol violation in randomization procedures, 30 participants (6 in reasoning training arm) were dropped from the analyses. This resulted in a total sample of 2,802 persons with 699 assigned to reasoning training.

Participants received training according to standardized procedures across 10 sessions over 5- to 6-week period. Regardless of training group assignment, all three interventions shared common design features: (1) equivalent intervention dosing, (2) small group settings with individual and group exercises by certified trainers, (3) ten 60 to 75 min sessions, (4) focus on training of strategies relevant to the ability trained, (5) modeling of strategy use by trainer and practice by participants, (6) feedback on performance, (7) fostering self-efficacy regarding performance, (8) applying learned strategies to real-world tasks, and (9) social interaction. In reasoning training, the first five sessions focused on strategy instruction and exercises to practice the strategy, while the last five sessions provided additional practice exercises, but no new strategy training. Training was administered over six replicates or waves.

Immediately following training (<10 days), participants completed a post-test assessment including three sessions similar to the baseline assessment. Follow-up assessments occurred again at approximately 1, 2, 3, and 5 years following training.

Just prior to the first annual assessment, booster training was offered to a 60% random subsample of compliant participants, who had completed at least 8 out of 10 initial training sessions. Booster training consisted of four 75-min refresher classes designed to reinforce strategies learned during the initial training sessions to help participants enhance and maintain training gains; no new information was provided during booster sessions. Booster training was again offered to this same subsample just prior to the third annual assessment.

Measures

Dependent measures—Cognitive outcome analysis. The dependent variables in this analysis were three reasoning measures and a composite score of the three measures. The *Letter Series* test requires participants to identify the pattern in a series of letters and circle the letter that comes next in the series (30 items; Time: 6 min; Thurstone & Thurstone, 1949). The *Word Series* test requires participants to identify the pattern in a series of words, such as the month or day of the week, and circle the word that comes next in the series (30 items; Time: 6 min; Gonda & Schaie, 1985). The *Letter Sets* test requires participants to identify which set of letters out of four letter sets does not follow the pattern of letters (15 items; 6 min; Ekstrom, French, Harman, & Derman, 1976). The scores for each test were transformed using the Blom transformation to reduce skewness (Blom, 1958). For the *Reasoning Composite*, each of the three reasoning measures was standardized to its baseline value, and an average of the equally weighted standardized scores was calculated.

Dependent measures—Functional outcome analysis. The dependent variables in this analysis were two measures of everyday reasoning/problem-solving abilities—the Everyday Problems Test (EPT; Marsiske & Willis, 1993) and the Observed Tasks of Daily Living (OTDL; Diehl et al., 2005)—and two measures of everyday speed of processing—the Complex Reaction Time test (CRT; Ball, 2000) and the Timed Instrumental Activities of Daily Living (TIADL; Owsley, Sloane, McGwin, & Ball, 2002). Lower scores on the CRT and TIADL reflected better performance. The scores for each test were transformed using the Blom transformation to reduce skewness (Blom, 1958).

Covariates. The covariates were baseline MMSE, self-rated health, age, education, and gender. The MMSE (Folstein, Folstein, & McHugh, 1975) assesses orientation, attention, language, memory, and construction skills and is widely used as a measure of general mental status (Score = 0-30). Individuals with scores below 23 were excluded at baseline screening, resulting in a truncated range from 23 to 30 ($M = 27.26$, $SD = 2.02$). Participants completed the self-rated health item ("in general, would you say your health is") of the MOS 36-item short-form health survey (SF-36) (Ware & Sherbourne, 1992) with a 5-point scale (1 = *excellent*, 2 = *very good*, 3 = *good*, 4 = *fair*, 5 = *poor*). For the analysis, health was coded dichotomously as 0 = *excellent/very good/good*, 1 = *fair/poor*, given the sample distribution. Gender was coded as 0 = male, 1 = female. Age, education, and MMSE were mean-centered.

Adherence indicators. Participants were considered compliant with initial training if they participated in at least 80% of the training sessions (i.e., 8-10 sessions). Adherence with the booster training sessions at the first annual and third annual follow-up assessments was indicated by participation in at least three of the four sessions; participants not randomly assigned to booster training were given missing values for the booster adherence variables. Given that 90% of participants were compliant with initial training and 98% were compliant with first annual and third annual boosters, adherence indicators were coded as 0 = compliant, 1 = noncompliant.

Reasoning Training Program

Reasoning 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; Schaie & Willis, 1986). Such problems involve identifying the pattern in a series of letters or words. Participants were taught strategies (e.g., underlining repeated letters, putting slashes between series, indicating skipped items in a series with tick marks) to identify the pattern or sequence involved in solving a problem; they used the pattern to determine the next item in the series. Participants practiced the strategies in individual and group exercises. Exercises involved abstract reasoning tasks (e.g., letter series) and reasoning problems related to ADLs (e.g., identifying medication dosing pattern).

Data Analysis

To examine the trajectory of performance on the outcome measures throughout the several phases of the ACTIVE training protocol that occurred between the baseline assessment and the fifth annual assessment, piecewise latent growth modeling was utilized (see, for example, Raudenbush & Bryk, 2002,

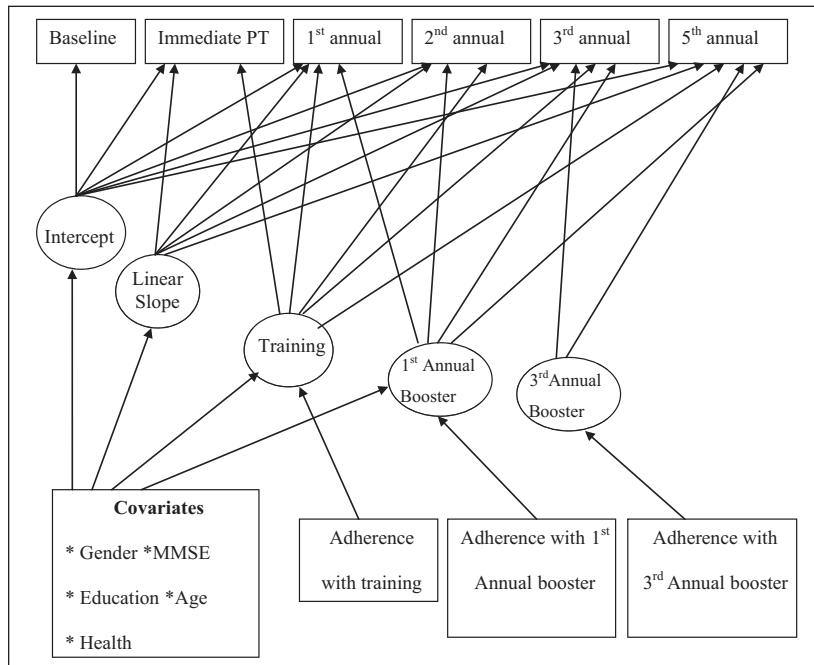


Figure 1. Piecewise growth model with demographic covariates and adherence indicators as predictors.

Note. MMSE = Mini-Mental State Exam.

pp. 178-179). All analyses were conducted using Amos 7 (Arbuckle, 2006). Full information maximum likelihood (FML) was used to estimate the model parameters to utilize all available data in the total sample, even in the presence of some missing values on the variables.

As shown in Figure 1, the growth model included an intercept and linear slope to indicate the average performance at baseline and linear change over time, *without the influence of training or booster effects*. All factor loadings from the intercept to the six observed time points were fixed to 1; for the linear slope factor, these paths were set to 0, 0.23, 1.23, 2.23, 3.23, and 5.23, representing amount of time in years since baseline. Thus, the mean linear slopes can be interpreted as the average rate of linear change over 1 year. Three additional latent factors were also included to represent (1) the influence of training, (2) the influence of the booster training received at the first annual assessment, and (3) the influence of booster training received at the third annual assessment. Each of these latent factors assessed the effect of that component of training and was operationalized in the model with

Table 2. Fit of the Unconditional and Conditional Piecewise Growth Models (N = 699).

| Outcome | Chi-square | df | p value | CFI | NFI | RMSEA |
|----------------------|------------|---------------------|---------|------|------|-------|
| Unconditional models | | | | | | |
| Letter series | 19.51 | 16 | .243 | 1.00 | 1.00 | .02 |
| Letter sets | 22.36 | 17 ^a | .171 | 1.00 | 1.00 | .02 |
| Word series | 34.40 | 16 | .005 | 1.00 | 0.99 | .04 |
| Reasoning composite | 27.70 | 16 | .035 | 1.00 | 1.00 | .03 |
| EPT | 25.98 | 18 ^{a,b} | .100 | 1.00 | 1.00 | .03 |
| OTDL | 12.89 | 8 ^a | .116 | 1.00 | 0.99 | .03 |
| TIADL | 18.25 | 19 ^{a,b,c} | .148 | 1.00 | 1.00 | .02 |
| CRT | 28.85 | 13 ^{a,b} | .007 | 1.00 | 0.99 | .04 |
| Conditional models | | | | | | |
| Letter series | 55.42 | 36 | .020 | 1.00 | 0.99 | .03 |
| Letter sets | 50.43 | 37 ^a | .069 | 1.00 | 0.99 | .02 |
| Word series | 65.86 | 36 | .002 | 0.99 | 0.98 | .03 |
| Reasoning Composite | 58.00 | 36 | .011 | 1.00 | 0.99 | .03 |
| EPT | 48.76 | 38 ^a | .113 | 1.00 | 0.99 | .02 |
| OTDL | 30.86 | 27 ^{a,d} | .227 | 1.00 | 0.99 | .01 |
| TIADL | 56.32 | 39 ^{a,b,c} | .036 | 0.99 | 0.98 | .03 |
| CRT | 53.54 | 33 ^a | .013 | 0.99 | 0.98 | .03 |

Note. Measurement errors were constrained to equality for all proximal outcomes, for EPT, and for TIADL. CFI = Comparative Fit Index; NFI = Normed Fit Index; RMSEA = root mean square error of approximation; EPT = Everyday Problems Test; OTDL = Observed Tasks of Daily Living; TIADL = Timed Instrumental Activities of Daily Living; CRT = Complex Reaction Time.

^aVariance for the booster effect at the third annual assessment was constrained to zero.

^bVariance for the training effect was constrained to zero.

^cVariance for the booster effect at the first annual assessment was constrained to zero.

^dVariance for the linear slope was constrained to zero.

loadings fixed to 0 for time points prior to the training component and loadings fixed to 1 for time points at or after that training component.

Unconditional (No-Predictor) models

As our first step in the analyses, unconditional models (i.e., without predictors) were estimated for each outcome measure. As noted in Table 2, several of the models had one or more negative error variance estimates, most likely reflecting difficulty in estimating near-zero interindividual variability for the particular effect. In all cases, these variances were constrained to zero to allow for an admissible solution. Next, homogeneity of the measurement

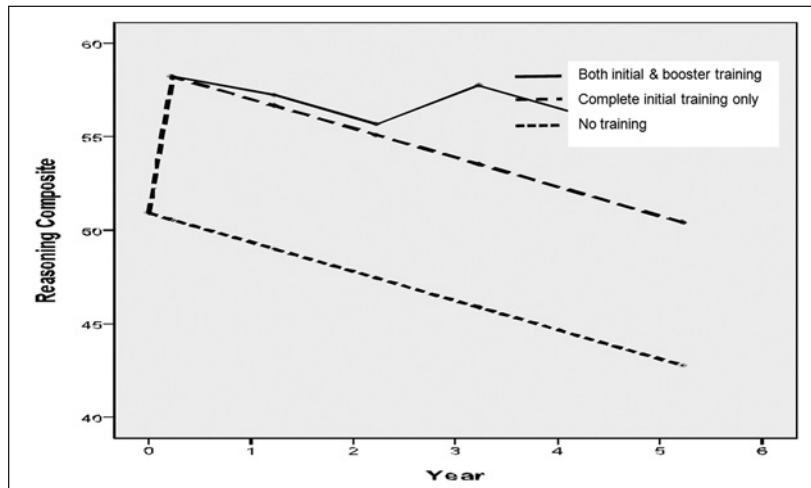


Figure 2. Reasoning Composite: Predicted trajectories of participants with no training, complete initial training only, and initial training and booster training.

errors across time for each outcome was tested and retained for all of the cognitive outcomes and two of the functional outcomes (EPT and TIADL). As shown in Table 2, the final unconditional models for all outcomes showed very good fit to the data.

Conditional Models: Influence of Adherence and Demographic Covariates

After finding a good fit for the unconditional models, the demographic covariates (i.e., age, education, MMSE, gender, and self-rated health) and three adherence indicators (training adherence, first annual booster adherence, and third annual booster adherence) were added to the models. These conditional models also showed good fit to the data, as shown in Table 2. The estimated means and variances for the piecewise growth model factors are given in Table 4. The three continuous covariates were mean-centered. Three predicted trajectories controlling for adherence and demographic variables are shown in Figure 2; specifically, the trajectory that would have been predicted without training is shown and is contrasted with the trajectories predicted for those who receive and complete only the initial training as well as those who had the initial training and the booster training. Outcome variables were transformed to *T*-scores, with a mean of 50 and standard deviation of 10 at baseline.

Results

Results are presented for the three major questions addressed. Descriptive statistics for the cognitive and functional outcome measures across the six time points assessed are shown in Table 3 for the booster and no booster groups.

1. What are the trajectories for the reasoning training group through the fifth annual follow-up on the cognitive and functional outcomes?

Cognitive (Reasoning) Outcomes

Significant linear decline was observed in all reasoning measures without training. However, reasoning training resulted in a significant, positive training effect for all reasoning measures; training effects were found at immediate posttest and were maintained through fifth annual follow-up (Table 3; Figure 2). Given that our dependent variables were scaled in *T*-score metric ($M = 50$, $SD = 10$), the average training gains were equivalent to approximately 0.8 *SD* for Letter Series, Word Series, and the Reasoning Composite and approximately 0.4 *SD* for Letter Sets.

Functional Outcomes

A significant training effect was observed for only one functional outcome, the CRT (Tables 3 and 4). Participants showed a significant decrease in reaction time on this measure. Given that our dependent variables were scaled in *T*-score metric ($M = 50$, $SD = 10$), the average training gain for CRT was equivalent to approximately 0.2 *SD*.

2. Does intervention dosing (i.e., adherence with training regimen and boosters) influence outcomes?

Adherence

With regard to cognitive (reasoning) outcomes, training adherence and booster adherence had a positive effect. Training adherence (80%+ sessions) resulted in greater training effects for Letter Series ($p < .001$), Word Series ($p < .001$), and the reasoning composite ($p < .001$). Moreover, booster adherence (75%+ sessions) resulted in a greater booster effect at the first annual booster for all cognitive outcomes.

Table 3. Descriptive Statistics for the Cognitive and Functional Outcomes Across Time by Assignment to Booster.

| Outcome | Group | Baseline | PT | 1st annual | 2nd annual | 3rd annual | 5th annual | |
|---------------------|------------|----------|-------|------------|------------|------------|------------|--|
| Letter series | Booster | | | | | | | |
| | <i>M</i> | 50.25 | 58.65 | 58.38 | 57.11 | 60.38 | 57.94 | |
| | <i>SD</i> | 10.40 | 10.26 | 10.29 | 10.09 | 10.42 | 9.94 | |
| | <i>n</i> | 370 | 355 | 299 | 288 | 274 | 246 | |
| | No booster | | | | | | | |
| | <i>M</i> | 49.78 | 57.96 | 56.04 | 55.67 | 57.03 | 56.72 | |
| Letter sets | Booster | | | | | | | |
| | <i>M</i> | 50.26 | 54.57 | 55.08 | 52.93 | 54.97 | 52.99 | |
| | <i>SD</i> | 10.15 | 10.36 | 10.13 | 10.63 | 9.80 | 10.27 | |
| | <i>n</i> | 368 | 355 | 298 | 285 | 273 | 245 | |
| | No booster | | | | | | | |
| | <i>M</i> | 49.76 | 53.52 | 51.92 | 52.20 | 53.04 | 52.19 | |
| Word series | Booster | | | | | | | |
| | <i>M</i> | 50.36 | 58.18 | 58.12 | 57.36 | 59.64 | 56.89 | |
| | <i>SD</i> | 9.84 | 9.99 | 10.27 | 10.35 | 10.71 | 10.64 | |
| | <i>n</i> | 370 | 353 | 304 | 283 | 276 | 246 | |
| | No booster | | | | | | | |
| | <i>M</i> | 49.60 | 56.95 | 55.98 | 55.92 | 56.84 | 54.80 | |
| Reasoning composite | Booster | | | | | | | |
| | <i>M</i> | 50.32 | 57.93 | 57.94 | 56.39 | 59.14 | 56.64 | |
| | <i>SD</i> | 10.15 | 10.46 | 10.36 | 10.52 | 10.48 | 10.54 | |
| | <i>n</i> | 370 | 357 | 305 | 290 | 280 | 246 | |
| | No booster | | | | | | | |
| | <i>M</i> | 49.67 | 56.77 | 55.12 | 55.02 | 56.21 | 55.03 | |
| EPT | Booster | | | | | | | |
| | <i>M</i> | 50.70 | 51.76 | 51.41 | 50.85 | 51.44 | 51.01 | |
| | <i>SD</i> | 10.19 | 10.19 | 10.12 | 10.45 | 10.14 | 10.56 | |
| | <i>n</i> | 370 | 354 | 316 | 309 | 288 | 270 | |
| | No booster | | | | | | | |
| | <i>M</i> | 49.18 | 50.33 | 50.37 | 50.30 | 51.14 | 50.67 | |
| | <i>SD</i> | 9.73 | 10.15 | 10.23 | 9.79 | 9.93 | 10.28 | |
| | <i>n</i> | 328 | 268 | 229 | 222 | 210 | 175 | |

(continued)

Table 3. (continued)

| Outcome | Group | Baseline | PT | 1st annual | 2nd annual | 3rd annual | 5th annual | |
|-----------|------------|----------|-------|------------|------------|------------|------------|--|
| OTDL | Booster | | | | | | | |
| | <i>M</i> | 50.08 | — | 53.17 | 53.20 | 53.77 | 53.97 | |
| | <i>SD</i> | 9.52 | — | 9.52 | 11.11 | 10.02 | 11.64 | |
| | <i>n</i> | 370 | — | 297 | 283 | 275 | 246 | |
| | No booster | | | | | | | |
| | <i>M</i> | 48.66 | — | 51.91 | 52.84 | 52.39 | 51.38 | |
| <i>SD</i> | 9.88 | — | 10.78 | 10.42 | 10.31 | 11.86 | | |
| <i>n</i> | 328 | — | 210 | 201 | 193 | 157 | | |
| TIADL | Booster | | | | | | | |
| | <i>M</i> | 49.65 | 48.34 | 48.73 | 48.90 | 49.46 | 50.01 | |
| | <i>SD</i> | 9.76 | 9.95 | 10.14 | 10.61 | 9.45 | 9.98 | |
| | <i>n</i> | 370 | 353 | 298 | 286 | 276 | 247 | |
| | No booster | | | | | | | |
| | <i>M</i> | 50.42 | 50.16 | 48.91 | 48.32 | 49.31 | 49.35 | |
| <i>SD</i> | 10.16 | 10.12 | 10.22 | 9.53 | 9.99 | 9.46 | | |
| <i>n</i> | 329 | 270 | 210 | 201 | 193 | 157 | | |

Note. PT = Post Test; EPT = Everyday Problems Test; OTDL = Observed Tasks of Daily Living; TIADL = Timed Instrumental Activities of Daily Living.

With regard to functional outcomes, only booster adherence had an effect. In contrast to the positive effects of adherence for the cognitive measures, training adherence did not have a positive effect for any of the functional outcomes. However, booster adherence resulted in greater booster effects at first annual booster for the OTDL ($p = .013$), TIADL ($p = .020$), and CRT ($p = .003$), but in poorer booster effects at the third annual assessment for EPT ($p = .013$).

Booster

Only Word Series showed a significant effect from the first annual booster. However, at the third annual booster, all cognitive (reasoning) outcomes showed a significant positive effect (Tables 3 and 4; Figure 2). Regarding functional outcomes, at the first annual booster, CRT and OTDL showed significant booster effects.

3. Are covariates significant predictors of training effects?

Table 4. Means and Variances for the Piecewise Growth Model Factors with Demographic Covariates and Adherence Indicators Included in the Model.

| Outcome | Intercept | | Linear slope | | Training effect | | 1st annual booster effect | | 3rd annual booster effect | |
|---------------------|-----------|----------|--------------|-------------------|-----------------|-------------------|---------------------------|-------------------|---------------------------|-------------------|
| | M | Variance | M | Variance | M | Variance | M | Variance | M | Variance |
| Letter series | 50.36*** | 43.50*** | -1.09*** | 0.07 | 8.05*** | 6.70*** | 0.07 | 0.91 | 3.36*** | 2.04 |
| Letter sets | 51.50*** | 41.27*** | -1.41*** | 0.05 | 4.43*** | 8.00*** | 0.19 | 1.29 | 3.07* | 0.00 ^a |
| Word series | 50.44*** | 44.67*** | -1.83*** | 0.39** | 7.94*** | 6.54*** | 1.41* | 0.36 | 2.96** | 1.95 |
| Reasoning composite | 50.92*** | 48.26*** | -1.56*** | 0.24** | 7.65*** | 5.24*** | 0.58 | 1.37 | 3.64*** | 1.65 |
| EPT | 51.18*** | 42.58*** | -0.54* | 0.33*** | 0.56 | 0.00 ^a | -0.36 | 1.55 | 0.61 | 0.00 ^a |
| OTDL | 48.89*** | 26.83*** | 0.04 | 0.00 ^a | — | — | 2.26* | 10.75*** | -0.32 | 0.00 ^a |
| TIADL | 51.61*** | 34.24*** | 0.87* | 0.09 | -1.35 | 0.00 ^a | -0.80 | 0.00 ^a | -0.63 | 0.00 ^a |
| CRT | 47.72*** | 62.44*** | 0.43 | 0.71*** | -1.76*** | 0.00 ^a | -1.59* | 4.64* | -0.55 | 0.00 ^a |

Note. EPT = Everyday Problems Test; OTDL = Observed Tasks of Daily Living; TIADL = Timed Instrumental Activities of Daily Living; CRT = Complex Reaction Time.

^aThis variance was constrained to zero.

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

Predictors of baseline performance level and predictors of slope (training effect) were examined.

Predictors of Cognitive Outcomes

Predictors of higher baseline performance for all reasoning measures were younger age, higher education, higher MMSE (at baseline), and better health. A higher MMSE at baseline was the only predictor of the training effect (slope; $p = .006$). In contrast, a higher baseline MMSE score was associated with smaller booster effects (slope; $p = .001$) at the first and third annual assessments. Finally, better health was related to a greater training effect (slope) for Letter Series ($p = .047$).

Predictors of Functional Outcomes

Predictors of higher baseline performance on all functional outcomes were younger age, higher MMSE (at baseline), and better health ($p < .001$). Moreover, higher education was associated with higher baseline performance for EPT, OTDL, and TIADL ($p < .001$), and being female was associated with better TIADL performance ($p < .001$) but lower CRT performance ($p = .005$). Younger age was also related to a slower linear decline (without influence of training or booster effects) in TIADL ($p = .013$) and a larger training effect (slope) for EPT ($p = .004$) and CRT ($p = .038$). Poorer health showed an association with greater linear decline for OTDL ($p = .021$).

Discussion

This study addressed three major questions with regard to the reasoning training group within the ACTIVE trial. What was the impact of training on the trajectory of the reasoning trained group from baseline to 5-year follow-up? Did adherence with training and booster sessions influence training outcomes? What covariates were significant predictors of training effects? We will first discuss these findings with regard to cognitive outcomes and then consider the more limited findings regarding functional outcomes.

The magnitude of age-related decline for reasoning ability over the 5-year interval (without training) was found to be approximately 0.78 *SD* units for the Reasoning Composite (Table 4). The magnitude of the initial training effect was on the order of approximately 0.80 *SD* units for the Reasoning Composite. Thus, the magnitude of benefit derived from the initial training (10 sessions) was on the order of the magnitude of age-related decline over the 5-year follow-up interval. The comparability of magnitude of training

gain from initial training and magnitude of age-related decline over 5 years may contribute to our understanding of the maintenance of training effects. The descriptive statistics for reasoning performance (Table 3) indicate that, on average, participants were performing at least 0.50 *SD* units above baseline even 5 years after training, for the Reasoning Composite.

We can offer two hypotheses about why there might have been a lesser effect for the Letter Sets measure. First, less time in initial training and in booster sessions was focused on Letter Sets; this allocation of training time was based in part on prior factor analysis of markers of Inductive Reasoning and the finding that Letter Sets is a less salient marker of Inductive Reasoning ability. Second, although the strategies focused on during training do apply to Letter Sets problems, the application of the strategies to Letter Set problems is more complex and thus strategy training may have been less useful in solving Letter Set items.

The second major question of the study focused on dosing issues. Although the efficacy of cognitive training has been demonstrated in prior research, these smaller scale studies have given less attention to issues of dosing. How much training is needed for a certain level of effect? Do temporally spaced booster sessions significantly increase the training effect? These dosing issues are relevant to cost-benefit analyses of training and for the pragmatics of delivery of cognitive interventions. Amount of training required for a given effect also has implications for the diversity of elderly participants (e.g., frailty) who may be offered training. The ACTIVE trial is unique in its design and focus on addressing these issues. The study findings suggest that adherence with training and added booster sessions can significantly enhance the training experience and increase the training effect. Training adherence for the initial training and for the booster sessions resulted in larger training effects for the Reasoning Composite and for the individual measures of Letter Series and Word Series. However, it should be noted that the majority of subjects were adherent to both initial and booster training; this high rate of compliance may be partly due to opportunities for make-up sessions and the rigorous monitoring of compliance by study personnel. That most subjects did adhere to study protocol may have limited the power to detect the effects of variability in adherence.

The findings regarding booster sessions indicate that a delayed booster, at 3 years after training, had significant added training benefit. The magnitude of the booster effect at third annual was approximately 0.30 *SD* units. The impact of the third annual booster was approximately one-half (48%) as great as the magnitude of initial training for the Reasoning Composite. The effect of the third annual booster was three times as large as the amount of age-related decline that would occur in 1 year with no training. Descriptive statistics,

shown in Table 3 (also Figure 2), indicate that the highest performance for the boosted group across the 5-year trajectory was immediately after the third Annual booster training. In contrast, there was less improvement associated with the booster immediately before the first annual follow-up, with the exception of Word Series. It may be that, on average, many of the training participants were functioning at a high level after initial training and less able to benefit from the booster that so closely followed the initial training.

Level of reasoning scores at baseline (intercept) is associated with better functioning on a number of the covariates (younger age, higher education, higher MMSE, and better health). However, of note is that neither age nor gender was significant predictors of initial training effects (slope) or booster effects for any reasoning measures, indicating the generality of training gain across gender and the wide age range of participants from 65 to 90 years of age. In contrast, those with lower levels of education exhibited greater training effects for the Letter Series and Word Series. Likewise, participants with lower baseline MMSE showed greater booster effects for the Reasoning Composite; this is particularly noteworthy for booster effects occurring 3 years after training. Thus, additional booster interventions appear particularly useful for the cognitively less advantaged. The general self-report measure of health had little relationship with training or booster effects.

More limited effects of training were found for functional outcomes. Only one functional outcome (CRT) showed a training effect associated with initial training and the magnitude of training gain was only one fourth ($0.2 SD$) the size of the cognitive training effect. CRT is a speed-based measure, while training focused on strategies for solving reasoning problems. The focus on strategies specific to reasoning may have contributed to reduced transfer to functional outcomes in two ways. First, a focus solely on reasoning per se may have reduced transfer; performance on reasoning accounted for a much smaller proportion of variability in performance on functional measures than on reasoning measures. Second, the focus on strategies may have emphasized to subjects the importance of accuracy over an emphasis on speed.

One of the more interesting findings is that both additional training (booster) and adherence to booster were associated with improvement on multiple functional measures (CRT, TIADL, OTDL) at third annual booster. Both the TIADL and OTDL focus on performance of instrumental tasks of daily living. The functional improvement associated with booster training and adherence may suggest that additional and more intensive training is needed to demonstrate a far transfer effect to tasks of daily living; the finding of the effect at third annual may suggest the importance of temporal spacing of interventions to allow for integration of cognitive effects. A younger age was the only predictor of training for functional outcomes.

In summary, the ACTIVE study is the first large-scale randomized trial to show that cognitive training improves cognitive functioning in well-functioning older adults, and that this improvement lasts up to 5 years follow-up. Prior smaller intervention studies had documented significant immediate effects of training; the ACTIVE trial using intent-to-treat analyses replicated these findings. However, prior training research had not carefully examined issues of adherence with training and the effect of temporally spaced booster sessions. Prior studies had seldom reported the proportion of participants compliant with the intervention or whether adherence enhanced the intervention effect. The significant effect of adherence indicates that the dosing of the intervention is an important factor in its effectiveness. The high proportion of participants who were compliant with training indicates that participants are willing and able to tolerate the intensity of intervention provided in ACTIVE. The finding that the 3-year booster sessions resulted in an effect approximately half the size of the initial training is informative, given that the number of booster sessions was 60% of the intensity of the initial training and the participants were 3 years older, on average in their mid-to-late 70s. The efficacy of the delayed booster suggests that maintenance of training effects may indeed extend beyond the 5-year follow-up, underscoring the importance of following this sample into old-old age.

Authors' Note

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Nursing Research, National Institute on Aging, or the National Institutes of Health. Representatives of the funding agency have been involved in the review of the manuscript but not directly involved in the collection, management, analysis, or interpretation of the data.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: ACTIVE is supported by grants from the National Institute on Aging and the National Institute of Nursing Research to Hebrew Senior Life (U01NR04507), Indiana University School of Medicine (U01NR04508), Johns Hopkins University (U01AG14260), New England Research Institutes (U01AG14282), Pennsylvania State University (U01AG14263), University of Alabama at Birmingham (U01AG14289), University of Florida (U01AG14276).

References

- Allaire, J. C., & Marsiske, M. (2002). Well- and ill-defined measures of everyday cognition: Relationship to older adults' intellectual ability and functional status. *Psychology and Aging, 17*, 101-115.
- Arbuckle, J. (2006). *Amos 7.0 user's guide*. Chicago, IL: SPSS.
- Ball, K. (2000). Increased mobility and reducing accidents of older drivers. In K. Schaie & M. Pietruha (Eds.), *Mobility and transportation in the elderly* (Vol. 5, pp. 2271-2281). New York, NY: Springer.
- Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., . . . Willis, S. L. (2002). Effects of cognitive training interventions with older adults: A randomized controlled trial. *Journal of the American Medical Association, 288*, 2271-2281.
- Ball, K., Edwards, J. D., & Ross, L. A. (2007). The impact of speed of processing training on cognitive and everyday functions. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences, 62*, 19-31.
- Blieszner, R., Willis, S., & Baltes, P. (1981). Training research in aging on the fluid ability of inductive reasoning. *Journal of Applied Developmental Psychology, 2*, 247-265.
- Blom, G. (1958). *Statistical estimates and transformed beta-variables*. New York, NY: Wiley.
- Boron, J. B., Willis, S. L., & Schaie, K. W. (2007). Cognitive training gain as a predictor of mental status. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences, 62*, P45-P52.
- Diehl, M., Marsiske, M., Horgas, A. L., Rosenberg, A., Saczynski, J. S., & Willis, S. L. (2005). The revised observed tasks of daily living: A performance-based assessment of everyday problem solving in older adults. *Journal of Applied Gerontology, 24*, 211-230.
- Diehl, M., Willis, S. L., & Schaie, K. W. (1995). Everyday problem solving in older adults: Observational assessment and cognitive correlates. *Psychology and Aging, 10*, 478-491.
- Ekstrom, R., French, J., Harman, H., & Derman, D. (1976). *Kit of Factor-Referenced Cognitive Tests* (Rev. ed.). Princeton, NJ: Educational Testing Service.
- 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*, 189-198.
- Gonda, J., & Schaie, K. (1985). *Schaie-Thurstone Mental Abilities Test: Word series test*. Palo Alto, CA: Consulting Psychologists Press.
- Jobe, J. B., Smith, D. M., Ball, K., Tennstedt, S. L., Marsiske, M., Willis, S. L., . . . Kleinman, K. (2001). ACTIVE: A cognitive intervention trial to promote independence in older adults. *Controlled Clinical Trials, 22*, 453-479.
- Kramer, A. F., & Willis, S. L. (2002). Enhancing the cognitive vitality of older adults. *Current Directions in Psychological Science, 11*, 173-177.
- Lezak, M. (1995). *Neuropsychological assessment*. Oxford, UK: Oxford University Press.

- Margrett, J., & Willis, S. (2006). In-home cognitive training with older married couples: Individual versus collaborative learning. *Aging, Neuropsychology, and Cognition: A Journal on Normal and Dysfunctional Development*, *13*, 173-195.
- Marsiske, M., & Willis, S.L. (1993). *Manual for the Everyday Problems Test*. University Park: The Pennsylvania State University.
- Marsiske, M., & Willis, S. L. (1995). Dimensionality of everyday problem solving in older adults. *Psychology and Aging*, *10*, 269-283.
- McArdle, J. J., & Prindle, J. J. (2008). A latent change score analysis of a randomized clinical trial in reasoning training. *Psychology and Aging*, *23*, 702-719.
- Owsley, C., Sloane, M., McGwin, G. Jr., & Ball, K. (2002). Timed instrumental activities of daily living tasks: Relationship to cognitive function and everyday performance assessments in older adults. *Gerontology*, *48*, 254-265.
- Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: SAGE.
- Saczynski, J., Margrett, J., & Willis, S. (2004). Older adults' strategic behavior: Effects of individual versus collaborative cognitive training. *Educational Gerontology*, *30*, 587-610.
- Saczynski, J., Willis, S., & Schaie, K. (2002). Strategy use in reasoning training with older adults. *Aging, Neuropsychology, and Cognition: A Journal on Normal and Dysfunctional Development*, *9*, 48-60.
- Salthouse, T. A., Fristoe, N., McGuthry, K. E., & Hambrick, D. Z. (1998). Relation of task switching to speed, age, and fluid intelligence. *Psychology and Aging*, *13*, 445-461.
- Schaie, K. (2005). *Developmental influences on adult intellectual development: The Seattle longitudinal study*. New York, NY: Oxford University Press.
- Schaie, K., & Willis, S. (1986). Can decline in adult intellectual functioning be reversed? *Developmental Psychology*, *22*, 223-232.
- Smith, G. E., Housen, P., Yaffe, K., Ruff, R., Kennison, R. F., Mahncke, H. W., & Zelinski, E. M. (2009). A cognitive training program based on principles of brain plasticity: Results from the Improvement in Memory with Plasticity-based Adaptive Cognitive Training (IMPACT) study. *Journal of the American Geriatrics Society*, *57*, 594-603.
- Stine-Morrow, E., & Basak, C. (2011). Cognitive intervention. In K. Schaie & S. Willis (Eds.), *Handbook of the psychology of aging* (7th ed., pp. 153-171). Amsterdam, Netherlands: Elsevier.
- Thurstone, L., & Thurstone, T. (1949). *Examiner manual for the SRA Primary Mental Abilities Test* (Form 10-14). Chicago, IL: Science Research Associates.
- Verhaeghen, P., Marcoen, A., & Goossens, L. (1992). Improving memory performance in the aged through mnemonic training: A meta-analytic study. *Psychology and Aging*, *7*, 242-251.
- Ware, J. E. Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care*, *30*, 473-483.
- Willis, S., & Nesselroade, C. (1990). Long term effects of fluid ability training in old-old age. *Developmental Psychology*, *26*, 905-910.

- Willis, S., & Schaie, K. (1994). Cognitive training in the normal elderly. In F. Forette, Y. Christen, & F. Boller (Eds.), *Plasticité cérébrale et stimulation cognitiv*. Paris, France: Fondation National de Gérontologie.
- Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., . . . Wright, E. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *Journal of the American Medical Association, 296*, 2805-2814.