Effects of Cognitive Training Interventions With Older Adults
A Randomized Controlled Trial

Karlene Ball, PhD
Daniel B. Berch, PhD
Karin F. Helmers, PhD
Jared B. Jobe, PhD
Mary D. Leveck, PhD
Michael Marsiske, PhD
John N. Morris, PhD
George W. Rebok, PhD
Sharon L. Tennstedt, PhD
Frederick W. Unverzagt, PhD
Sherry L. Willis, PhD
for the ACTIVE Study Group

Earliest half of community-dwelling persons aged 60 years and older express concern about declining mental abilities. Although there is substantial evidence that many cognitive abilities and processes are related to measures of functional status, need for care, and quality of life, few studies have addressed whether improving cognitive functions might have short- or long-term effects on activities related to living independently.

Context  Cognitive function in older adults is related to independent living and need for care. However, few studies have addressed whether improving cognitive functions might have short- or long-term effects on activities related to living independently.

Objective  To evaluate whether 3 cognitive training interventions improve mental abilities and daily functioning in older, independent-living adults.

Design  Randomized, controlled, single-blind trial with recruitment conducted from March 1998 to October 1999 and 2-year follow-up through December 2001.

Setting and Participants  Volunteer sample of 2832 persons aged 65 to 94 years recruited from senior housing, community centers, and hospital/clinics in 6 metropolitan areas in the United States.

Interventions  Participants were randomly assigned to 1 of 4 groups: 10-session group training for memory (verbal episodic memory; n = 711), or reasoning (ability to solve problems that follow a serial pattern; n = 705), or speed of processing (visual search and identification; n = 712); or a no-contact control group (n = 704). For the 3 treatment groups, 4-session booster training was offered to a 60% random sample 11 months later.

Main Outcome Measures  Cognitive function and cognitively demanding everyday functioning.

Results  Thirty participants were incorrectly randomized and were excluded from the analysis. Each intervention improved the targeted cognitive ability compared with baseline, durable to 2 years (P < .001 for all). Eighty-seven percent of speed-, 74% of reasoning-, and 26% of memory-trained participants demonstrated reliable cognitive improvement immediately after the intervention period. Booster training enhanced training gains in speed (P < .001) and reasoning (P < .001) interventions (speed booster, 92%; no booster, 68%; reasoning booster, 72%; no booster, 49%), which were maintained at 2-year follow-up (P < .001 for both). No training effects on everyday functioning were detected at 2 years.

Conclusions  Results support the effectiveness and durability of the cognitive training interventions in improving targeted cognitive abilities. Training effects were of a magnitude equivalent to the amount of decline expected in elderly persons without dementia over 7- to 14-year intervals. Because of minimal functional decline across all groups, longer follow-up is likely required to observe training effects on everyday function.

JAMA. 2002;288:2271-2281

©2002 American Medical Association. All rights reserved.

(Reprinted) JAMA, November 13, 2002—Vol 288, No. 18 2271
even in old age. Moreover, sustained engagement in cognitively stimulating activities has been found to impact neural structure in both older humans and rodents. Conversely, limited education has been found to be a risk factor for dementia. There is also a sizeable body of literature documenting that different types of cognitive training programs have large and durable effects on the cognitive functioning of older adults, even in advanced old age. At the same time, several important issues remain understudied. First, prior cognitive training studies with older adults have often paid relatively little attention to the use of appropriate control groups, the representativeness or heterogeneity of participants, the generalizability of training findings beyond particular laboratories, or adherence of participants to training protocols. For example, it has not been uncommon for such studies to analyze only compliant participants. Second, the broader implications of training on daily functioning in older adults, for the most part, have not been studied.

The primary objective of the ACTIVE (Advanced Cognitive Training for Independent and Vital Elderly) trial was to test the effectiveness and durability of 3 distinct cognitive interventions in improving the performance of elderly persons on basic measures of cognition and on measures of cognitively demanding daily activities (eg, food preparation, driving, medication use, financial management). These interventions previously had been found successful in improving cognitive abilities under laboratory or small-scale field conditions. We hypothesized that the effects of cognitive training on primary outcomes will be largely mediated through the basic cognitive abilities being trained. The detailed hypotheses may be summarized by 2 points: each training group will perform better than the other training and control groups on their respective primary and proximal outcomes, and those groups that received booster training will perform better than those that did not receive booster training on their primary and proximal outcomes.

**METHODS**

**Participants**

The recruitment goal for the ACTIVE trial was to enroll a diverse sample of older adults who, at enrollment, were living independently in good functional and cognitive status. Recruitment was conducted from March 1998 to October 1999; 2-year follow-up data were collected through December 2001. Details of the recruitment procedures have been published elsewhere. Participants aged 65 to 94 years were enrolled across 6 field sites using a variety of sampling frames and recruitment strategies (state driver’s license and identification card registries, medical clinic rosters, senior center and community organization rosters, senior housing sites, local churches, and rosters of assistance and service programs for low-income elderly persons). Oral assent was obtained for brief telephone screening, and written informed consent was obtained in person from each potential participant prior to administration of in-person screening measures.

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]); had a self-reported diagnosis of Alzheimer 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 (eg, 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, or measured vision worse than 20/70 with best correction), hearing (interviewer-rated), or communicative ability (interviewer-rated) that would sufficiently impair performance to make participation impossible.

**Study Design**

The study protocol was approved by the institutional review boards at the University of Alabama at Birmingham; Wayne State University, Detroit, Mich; the Hebrew Rehabilitation Center for the Aged, Roslindale, Mass; the Johns Hopkins University School of Medicine, Baltimore, Md; Indiana University, Bloomington; Purdue University, Indianapolis, Ind; Pennsylvania State University, University Park; the University of Florida, Gainesville; and the New England Research Institutes, Watertown, Mass.

The ACTIVE trial was sponsored by the National Institute on Aging and the National Institute of Nursing Research, and was randomized, controlled, and single-blind, using a 4-group design, including a no-contact control group and 3 intervention groups (memory training, reasoning training, or speed-of-processing training). These 3 interventions were selected because they showed the most promise in smaller laboratory studies and had been related to instrumental activities of daily living (IADL). Each intervention group received a 10-session intervention, conducted by certified trainers, for 1 of 3 cognitive abilities—memory, inductive reasoning, or speed of processing. Assessors were blinded to participant intervention assignment. Training exposure and social contact were standardized across interventions so that each intervention served as a contact control for the other 2 interventions. Booster training was provided to a random subsample in each intervention group. Measurement points consisted of baseline tests, an immediate posttest (following the intervention), and 1 and 2 annual posttests.

**Interventions**

The interventions were conducted in small group settings in ten 60- to 75-
minute sessions over 5- to 6-week periods. These were behavioral interventions with no pharmacological component. In all 3 conditions, sessions 1 through 5 focused on strategy instruction and individual and group exercises to practice the strategy. Sessions 6 through 10 provided additional practice exercises but introduced no new strategies.

Memory training focused on verbal episodic memory. Participants were taught mnemonic strategies for remembering word lists and sequences of items, text material, and main ideas and details of stories. Participants received instruction in a strategy or mnemonic rule, exercises, individual and group feedback on performance, and a practice test. For example, participants were instructed how to organize word lists into meaningful categories and to form visual images and mental associations to recall words and texts. The exercises involved laboratory-like memory tasks (eg, recalling a list of nouns, recalling a paragraph), as well as memory tasks related to cognitive activities of everyday life (eg, recalling a shopping list, recalling the details of a prescription label).

Reasoning training focused on the ability to solve problems that follow a serial pattern. Such problems involve identifying the pattern in a letter or number series or understanding the pattern in an everyday activity such as prescription drug dosing or travel schedules. Participants were taught strategies to identify a pattern and were given an opportunity to practice the strategies in both individual and group exercises. The exercises involved abstract reasoning tasks (eg, letter series) as well as reasoning problems related to activities of daily living.

Speed-of-processing training focused on visual search skills and the ability to identify and locate visual information quickly in a divided-attention format. Participants practiced increasingly complex speed tasks on a computer. Task difficulty was manipulated by decreasing the duration of the stimuli, adding either visual or auditory distraction, increasing the number of tasks to be performed concurrently, or presenting targets over a wider spatial expanse. Difficulty was increased each time a participant achieved criterion performance on a particular task.

Eleven months after the initial training was provided, booster training was offered to a randomly selected 60% of initially trained subjects in each of the 3 intervention groups. Booster training was delivered in four 75-minute sessions over a 2- to 3-week period.

**Measures**

The ACTIVE trial had multiple outcomes, both proximal (cognitive abilities) and primary (daily function) (Table 1). Composites were created to represent each domain. Each composite was the average of 2 or 3 test scores, equally weighted, and was designed as a measure of ability rather than performance on a specific test.

Proximal outcomes permitted a test of the impact of the 3 interventions on the appropriate cognitive abilities. Memory assessment focused on episodic verbal memory tasks. Reasoning assessment focused on tasks requiring identification of patterns in letter or word series problems. Speed-of-processing assessment focused on identifying the minimum stimulus duration at which participants could identify and localize information, with 75% accuracy, under varying levels of cognitive demand.

Primary outcomes were aspects of functional activities, both performance-based and self-reported. Everyday problem solving represented the ability to reason and correctly identify information in common everyday stimuli (eg, medication labels, charts, forms). This was measured via paper-and-pencil testing and behavioral simulations of everyday tasks. Everyday speed emphasized the speed with which participants interacted with real-world stimuli. Participants were asked to look up a specific telephone number, find food items on a crowded shelf of groceries, find ingredients on food labels, count out specified amounts of change, find specific information on medicine bottles, and respond appropriately to different traffic signs. Activities of daily living (ADL) and instrumental activities of daily living included self-ratings drawn from the Minimum Data Set—Home Care (MDS-HC). Driving habits included self-ratings of driving difficulty and avoidance of specific driving situations.

Tests were standardized by pooling scores at all time points and applying a Blom transformation, producing more normally distributed scores. Scores for tests at each time point were standardized to the baseline mean and SD. If 1 or more tests of a composite were missing, the composite score was calculated as the average of the nonmissing tests.

**Analysis**

To evaluate the effects of ACTIVE training over 2 years, a repeated-measures, mixed-effects model was used. The dependent variables were the proximal and primary composites measured at 4 time points: baseline, immediate posttest, first annual evaluation (A1), and second annual evaluation (A2). At posttest only the cognitive variables, the Everyday Problems Test, and the primary speed composite were measured. The independent variables were restricted to the basic design features: fixed effects for training group (memory, reasoning, speed, control); time (3 or 4 points); booster training; field site; and replicate within site. Three interaction terms were chosen for their importance and interpretability: time \( \times \) training, representing the net effect of the trial; time \( \times \) booster, representing nonspecific effects of the additional social contact of attending booster training, regardless of content; and time \( \times \) booster \( \times \) training, representing the training-specific effects of each booster intervention. For this analysis, the repeated-measures model was fitted to the available data, ignoring missing data. Then, to determine if selective attrition influenced the trial results, missing data were imputed using multiple imputation.
procedures, and the analysis was repeated.

Hypotheses were tested by comparing outcome composite scores at later times (posttest, A1, and A2) to baseline scores and to control group scores, yielding net differences. The net effect of training at any time was defined as: (trained mean – control mean at later time) – (trained mean – control mean at baseline). Similarly, the net effect of each booster training was defined as: (booster mean – unboosted mean at later time) – (booster mean – unboosted mean at baseline). Results are expressed as effect sizes (ie, difference in means divided by intrasubject SD) to allow direct comparison of different outcomes. In addition, covariate-adjusted training effects were examined, with covariates of age, sex, cognitive status (MMSE score), years of education, and visual acuity. Given the substantial variation associated with field site and replicate, these 2 factors were also included as covariates in all analyses.

Secondary analyses investigated the percentage of participants who showed reliable improvement in each training group. A participant was classified as having improved reliably on a particular measure if his or her performance at a follow-up occasion exceeded baseline performance on that measure by 1 SEM. The formula for reliable change was computed as outlined by Dudek, and analyses were conducted using SAS v8.2 (SAS Institute Inc, Cary, NC).

RESULTS

Five thousand individuals were contacted for participation (FIGURE 1). A total of 2832 persons were eligible, 905 (18.1%) were ineligible, and 1263 (25.3%) refused (either directly or passively by not coming to any appointments) prior to randomization. Reasons for ineligibility were: cognitive impairment on the MMSE (270 [29.8%]), vision impairment (192 [21.2%]), unavailability due to schedule (202 [22.3%]), too young (85 [9.4%]), medical conditions predisposing to imminent decline or short life expectancy (79 [8.7%]), significant ADL disability (48 [5.3%]), impaired communication (15 [1.7%]), diagnosed Alzheimer disease (7 [0.8%]), and prior participation in cognitive training trials (7 [0.8%]). Enrollment at the field sites ranged from 405 to 498 participants. Thirty eligible persons were randomized inappropriately, thus violating protocol, and were excluded from analyses. The analytic sample consists of 2802 participants randomized by the New England Research Institutes with a concealed system. Intention-to-treat analyses were used.

### Table 1. Outcome Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Time Point</th>
<th>Mode of Administration</th>
<th>Reliability*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Immediate Posttest</td>
<td>First and Second Annual Posttests</td>
</tr>
<tr>
<td>Memory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hopkins Verbal Learning Test31</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Auditory Verbal Learning Test32</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Rivermead Behavioral Memory Test33</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Reasoning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word series34</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Letter series35</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Letter sets36</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Speed of processing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Useful Field of View (tasks 2-4)23,25,37</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Primary Outcome Composites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Everyday problem solving</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Everyday Problems Test38</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Observed Tasks of Daily Living31,39</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Everyday speed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complex Reaction Time (2 tests)8</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Timed Instrumental Activities of Daily Living (IADL)40</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Activities of Daily Living and IADL functioning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimum Data Set–Home Care (3 scores)41</td>
<td>X</td>
<td>X</td>
<td>Survey</td>
</tr>
<tr>
<td>Driving habits</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty, avoidance, (space)‡42</td>
<td>X</td>
<td>X</td>
<td>Interview</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interview</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Survey</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interview</td>
</tr>
</tbody>
</table>

*Test-retest correlations except where noted.
†Cronbach α.
‡“Space” indicates a driver’s extent of travel in a specified period of time.
§α or χ from reference cited.
Ineligible participants were comparable with eligible participants in age (mean, 77 years) and proportion of women (77%). Ineligible participants tended to have a higher percentage of nonwhite persons (48%) and lower cognitive function (mean MMSE score, 20.9). Potential participants who were eligible and randomized (n=2802) were comparable with the group that was eligible and not randomized (n=1263). Compared with the randomized group, the randomized group was more educated (13.5 vs 12.3 years), slightly younger (mean, 74 vs 75 years), and 80% of the sample was retained at the 2-year follow-up, despite the advanced age of the cohort. The net effect of ACTIVE training on the proximal (cognitive) composites is displayed in the lower section of Table 3. Each training program produced an immediate effect on its corresponding cognitive ability. It is important to note that while these analyses were conducted on Blom-transformed variables, a near-identical pattern of findings was obtained with the untransformed variables. Temporal trends in the mean cognitive composite scores by intervention group are shown in Figure 2.

The net effect of ACTIVE training on functional outcome composites is detailed in the lower section of Table 3.

**Figure 1. Flow of Participants Through the Trial**

```
5000 Individuals Assessed for Eligibility

2832 Randomized

711 Assigned to Receive Memory Training
  64 Had No Training
  27 Had Partial Training
  620 Completed Intervention

705 Assigned to Receive Reasoning Training
  33 Had No Training
  45 Had Partial Training
  627 Completed Intervention

712 Assigned to Receive Speed Training
  40 Had No Training
  35 Had Partial Training
  637 Completed Intervention

704 Assigned to Control

34 Deactivated
  1 Died
  25 Withdraw
  8 Site Decision
  31 Not Assessed
  639 Assessed

2168 Excluded

905 Ineligible
  1263 Refused Participation

711 Assigned to Receive Booster
  82 Did Not Receive Booster
  7 Received Partial Booster
  283 Completed Booster

710 Assigned to Receive Reasoning Booster
  66 Did Not Receive Booster
  4 Received Partial Booster
  301 Completed Booster

705 Assigned to Receive Speed Booster
  71 Did Not Receive Booster
  4 Received Partial Booster
  295 Completed Booster

372 Assigned to Receive Booster
  38 Deactivated
  10 Died
  25 Withdrawn
  3 Site Decision
  45 Not Assessed
  585 Assessed

371 Assigned to Receive Reasoning Booster
  37 Deactivated
  15 Died
  20 Withdrawn
  2 Site Decision
  56 Not Assessed
  566 Assessed

370 Assigned to Receive Speed Booster
  29 Deactivated
  11 Died
  14 Withdrawn
  4 Site Decision
  52 Not Assessed
  601 Assessed

37 Deactivated
  9 Died
  24 Withdrawn
  4 Site Decision
  49 Not Assessed
  584 Assessed

8 Protocol Violations
  703 Included in Analysis

6 Protocol Violations
  699 Included in Analysis

10 Protocol Violations
  702 Included in Analysis

6 Protocol Violations
  698 Included in Analysis
```

"Withdraw” indicates subjects withdraw for reasons including scheduling conflicts, poor health, and lack of interest in continuing; “site decision,” that subjects were withdrawn by study sites because they repeatedly missed appointments or were uncooperative or disruptive during testing sessions.

©2002 American Medical Association. All rights reserved. (Reprinted) JAMA, November 13, 2002—Vol 288, No. 18 2275
These effects were generally small on the effect-size scale—most were below 0.10—and did not differ significantly from zero at A1 or A2. It is important to note, however, that the vast majority of this sample remained functionally independent over the course of the 24-month observation period. For the crucial measures of ADL performance—measures that have been shown to predict movement into home care and institutional programs—a relatively low ADL decline rate (defined as ≥2 points on the summary measure) of 6% was observed at 12 months, with a modest increase to 8% at 24 months.

The impact of booster training at A1 and A2 is detailed in Table 4. Again, the strongest effects were seen in cognitive

---

**Table 2. Baseline Characteristics of Participants (N = 2802)**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Sample</th>
<th>General Population†</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD) [range], y</td>
<td>73.6 (5.9) [65-94]</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Age groups, y, %48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65-74</td>
<td>60.1</td>
<td>57.6</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>75-84</td>
<td>35.0</td>
<td>32.5</td>
<td></td>
</tr>
<tr>
<td>≥85</td>
<td>4.9</td>
<td>9.9</td>
<td></td>
</tr>
<tr>
<td>Women, %49</td>
<td>75.9</td>
<td>57.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Race, %49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>73.3</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>26.0</td>
<td>8.6</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Other/unknown</td>
<td>0.7</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>High school graduate, %49</td>
<td>88.6</td>
<td>67.0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Married, %49</td>
<td>35.9</td>
<td>56.6</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>MMSE score, mean (SD) [range]</td>
<td>27.3 (2.0) [23-30]</td>
<td>26.4 (2.1)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>SF-36 physical function, mean (SD)</td>
<td>68.8 (24.1) [0-100]</td>
<td>62</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*NA indicates not applicable; MMSE, Mini-Mental State Examination; and SF-36, Short Form Health Survey (36 item).
†Data from references in “characteristic” column.
‡For general population, only percentage given.
§For self-reported drivers only.
¶For general population, only percentage given.

**Table 3. Training Effects on Proximal and Primary Outcomes**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Memory Training</th>
<th>Reasoning Training</th>
<th>Speed Training</th>
<th>Control†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Net Effect Size (P Value)‡</td>
<td>Showing Reliable Improvement, %§</td>
<td>Net Effect Size (P Value)‡</td>
<td>Showing Reliable Improvement, %§</td>
</tr>
<tr>
<td>Proximal Outcome Composites</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memory Posttest</td>
<td>0.257 (&lt;.001)</td>
<td>26</td>
<td>−0.009</td>
<td>17</td>
</tr>
<tr>
<td>A1</td>
<td>0.212 (&lt;.001)</td>
<td>22</td>
<td>−0.011</td>
<td>11</td>
</tr>
<tr>
<td>A2</td>
<td>0.174 (&lt;.001)</td>
<td>40</td>
<td>−0.03</td>
<td>27</td>
</tr>
<tr>
<td>Reasoning Posttest</td>
<td>−0.018</td>
<td>34</td>
<td>0.480 (&lt;.001)</td>
<td>74</td>
</tr>
<tr>
<td>A1</td>
<td>0.021</td>
<td>34</td>
<td>0.402 (&lt;.001)</td>
<td>63</td>
</tr>
<tr>
<td>A2</td>
<td>0.045</td>
<td>36</td>
<td>0.257 (&lt;.001)</td>
<td>53</td>
</tr>
<tr>
<td>Speed</td>
<td>Posttest</td>
<td>−0.045</td>
<td>34</td>
<td>0.003</td>
</tr>
<tr>
<td>A1</td>
<td>−0.054</td>
<td>35</td>
<td>−0.033</td>
<td>34</td>
</tr>
<tr>
<td>A2</td>
<td>−0.034</td>
<td>36</td>
<td>−0.043</td>
<td>35</td>
</tr>
<tr>
<td>Primary Outcome Composites</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Everyday problem solving</td>
<td>−0.045</td>
<td>19</td>
<td>0.03</td>
<td>23</td>
</tr>
<tr>
<td>A1</td>
<td>−0.073</td>
<td>21</td>
<td>−0.027</td>
<td>25</td>
</tr>
<tr>
<td>A2</td>
<td>0.017</td>
<td>17</td>
<td>−0.125</td>
<td>18</td>
</tr>
<tr>
<td>ADL and IADL functioning</td>
<td>A1</td>
<td>0.02</td>
<td>17</td>
<td>−0.056</td>
</tr>
<tr>
<td>A2</td>
<td>0.017</td>
<td>17</td>
<td>−0.125</td>
<td>18</td>
</tr>
<tr>
<td>Everyday speed</td>
<td>Posttest</td>
<td>−0.091 (.02)</td>
<td>31</td>
<td>0.004</td>
</tr>
<tr>
<td>A1</td>
<td>−0.041</td>
<td>31</td>
<td>0.05</td>
<td>28</td>
</tr>
<tr>
<td>A2</td>
<td>−0.007</td>
<td>33</td>
<td>0.03</td>
<td>29</td>
</tr>
<tr>
<td>Driving habits</td>
<td>§</td>
<td>0.017</td>
<td>19</td>
<td>−0.052</td>
</tr>
<tr>
<td>A2</td>
<td>0.085</td>
<td>16</td>
<td>0.079</td>
<td>16</td>
</tr>
</tbody>
</table>

*Only significant P values reported. A1 indicates first annual evaluation; A2, second annual evaluation; ADL, activities of daily living; and IADL, instrumental activities of daily living.
†Net effect of the control is 0 at all time points, since net effect of group is defined as (group mean−control mean at time point)−(group mean−control mean at baseline).
‡Net difference divided by intra(subject SD (see “Methods” section).
§Calculated as the percentage of participants in each group who were ≥1 SEM above baseline.
¶Favorable response is in the negative direction.
‖For self-reported drivers only.
outcomes, where boosters for reasoning and speed training administered shortly before A1 produced significantly better performance. The impact of reasoning and speed booster training was greater at A1 than at A2. No effect was detected for memory booster on the memory proximal composite. Compared with those who did not receive booster training, participants randomized to speed booster performed significantly better at A1 on the functioning and everyday speed composites (P < .05), and marginally better at A2 (P < .10). Similarly, compared with those who did not receive booster training, participants randomized to reasoning booster performed marginally better on the functioning composite at A1 (P < .10).

The results of covariate-adjusted analyses were generally similar. While effect sizes were universally higher after adjusting for age, sex, education, visual function, and mental status, the overall pattern of results was the same and is not presented here. Similarly, analyses of imputed data sets did not differ in outcomes, suggesting that the trial results were not influenced by selective attrition.

Consistent with results of the primary analyses, secondary analyses indicated large immediate intervention gains on the cognitive outcomes. Eighty-seven percent of speed-trained, 74% of reasoning-trained, and 26% of memory-trained participants demonstrated reliable improvement on the pertinent cognitive composite immediately following intervention. While intervention participants showed reliable posttest gains, a comparable proportion of control participants also improved, and the proportion of control participants exhibiting reliable retest gain remained fairly constant across study intervals.

In terms of the proportion of the intervention group showing reliable gain in the trained domain, booster effects occurred for the speed conditions (boost, 92%; no boost, 68%; control, 32%) and the reasoning conditions (boost, 72%; no boost, 49%; control, 31%). While some dissipation of intervention effects occurred across time, cognitive effects were maintained from baseline to A2, particularly for boosted participants (79% [speed boost] vs 37% [controls]; 57% [reasoning boost] vs 35% [controls]).

**COMMENT**

To date, ACTIVE is the largest trial (N=2802) of cognitive interventions for the improvement of older adults’ performance on specific cognitive and per-

---

**Table 4. Net Effect of Booster Training on Proximal and Primary Outcomes**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Memory Training</th>
<th>Reasoning Training</th>
<th>Speed Training</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Net Effect Size (P Value)§</td>
<td>Showing Reliable Improvement, %‡</td>
<td>Net Effect Size (P Value)§</td>
</tr>
<tr>
<td></td>
<td>No Booster</td>
<td>Booster</td>
<td>No Booster</td>
</tr>
<tr>
<td>Proximal Outcome Composites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memory A1</td>
<td>0.044</td>
<td>23</td>
<td>21</td>
</tr>
<tr>
<td>A2</td>
<td>0.060</td>
<td>39</td>
<td>40</td>
</tr>
<tr>
<td>Reasoning A1</td>
<td>−0.009</td>
<td>35</td>
<td>32</td>
</tr>
<tr>
<td>A2</td>
<td>−0.036</td>
<td>37</td>
<td>35</td>
</tr>
<tr>
<td>Speed‡ A1</td>
<td>−0.03</td>
<td>37</td>
<td>33</td>
</tr>
<tr>
<td>A2</td>
<td>0.02</td>
<td>35</td>
<td>38</td>
</tr>
<tr>
<td>Primary Outcome Composites</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Everyday problem solving A1</td>
<td>−0.007</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>A2</td>
<td>−0.033</td>
<td>19</td>
<td>23</td>
</tr>
<tr>
<td>ADL and IADL functioning‡ A1</td>
<td>−0.088</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>A2</td>
<td>0.096</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>Everyday speed‡ A1</td>
<td>0.041</td>
<td>31</td>
<td>30</td>
</tr>
<tr>
<td>A2</td>
<td>−0.033</td>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>Driving habits‡† A1</td>
<td>−0.082</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>A2</td>
<td>−0.025</td>
<td>14</td>
<td>18</td>
</tr>
</tbody>
</table>

*Only significant P values reported. A1 indicates first annual evaluation; A2, second annual evaluation; ADL, activities of daily living; and IADL, instrumental activities of daily living.
†Net effect of the control is 0 at all time points, since net effect of group is defined as (group mean−control mean at time point)−(group mean−control mean at baseline).
§Net effect divided by intrasubject SD (see “Methods” section).
¶Calculated as the percentage of participants in each group who were ≥1 SEM above baseline.
|©2002 American Medical Association. All rights reserved. (Reprinted) JAMA, November 13, 2002—Vol 288, No. 18 2277 |
Overall, this large-scale study demonstrated that cognitive interventions helped normal elderly individuals to perform better on multiple measures of the specific cognitive ability for which they were trained. It did not, however, demonstrate the generalization of such interventions to everyday performance, at least in the initial 2 years. The effect sizes for the cognitive abilities at immediate posttest are for the most part consistent with previous research. Moreover, these effect sizes are comparable with or greater than the amount of longitudinal decline that has been reported in previous studies (TABLE 5), suggesting that these interventions have the potential to reverse age-related decline. Specifically, age-related decline for reasoning ability in samples of elderly persons without dementia has been found to be on the order of 0.22 SD over a 7-year interval (ages 67-74 years) and to increase to 0.42 SD over a 14-year interval (ages 67-81 years). Thus, immediate reasoning training effects (0.48 SD) were comparable with the amount of decline reported to occur in elderly persons without dementia over a 14-year interval. Likewise, decline in memory ability has been reported to be approximately 0.25 SD over a 6- or 7-year interval. Thus, memory training effects (0.25 SD) were comparable with the expected decline over a 7-year interval in elderly persons without dementia. Finally, decline for speed has been reported to be approximately 0.16 SD over 2 years. Immediate speed training effects (1.46 SD) were therefore 9 times greater than the expected decline over a 2-year interval in elderly persons without dementia.

Although training impact on the proximal composites decreased over time, it remained statistically significant, attesting to the durability of the intervention training effects. This is an important finding, since prior interventions (especially memory) have not shown 2-year durability. Furthermore, a very high percentage of trained participants achieved reliable improvement on the cognitive abilities, and ceiling effects at baseline on the cognitive measures explain lack of reliable improvement for most others. Of further note, the tests of training effects were conservative compared with those used in prior cognitive aging research. That is, prior cognitive training research has not used intention-to-treat analyses, instead excluding participants who dropped out or were noncompliant. In addition, prior research has not used diverse samples in terms of education and ethnicity. Thus, relative to prior work, training effects on cognitive abilities in this study are strong.

Insufficient sample size was ruled out as an explanation for the small effect sizes to date on the functional outcomes. The study was sufficiently powered to detect an effect size of 0.20 at 95% power with a sample of 2832. Power calculations were based on 6 Bonferroni-corrected, 2-sided comparisons with an overall α level of .05, a correlation of 0.7 between baseline and follow-up (based on pilot data), and an 80% completion rate. Based on these same assumptions, there was 90% power to detect booster training effects. Given that we retained more than 80% of the initial sample over the 2-year follow-up period and found no differential loss across treatment and control conditions, there should have been sufficient power to detect a significant effect of the cognitive training on the functional outcomes.

The absence of transfer to real-world outcomes is not particularly surprising. In addition to several decades of cognitive science research demonstrating the difficulty of obtaining such transfer, most of our subjects were not yet impaired in the domains of training. Indeed, there are several other potential explanations for the observed lack of transfer to daily function: the proportion of participants functioning at ceiling levels (ie, 43% had no room for improvement, as indicated by baseline performance within 1 SEM of the “best” value) on the daily functional composite, the evidence of strong practice or retest effects in the control group, and the control group’s lack of functional decline over the 2-year follow-up period.
In terms of the observed lack of functional decline in the control group, it is important to note that individuals with extant functional or cognitive decline were carefully screened out, and the study focused instead on intact individuals whose future decline rates were likely to mimic or be less than rates for the general elderly population. It was therefore unclear whether participants would show evidence of decline similar to established population parameters over the 12- and 24-month observation periods, or whether individuals in the sample would be more resilient and less subject to decline over such a short period of time. In specifying expected effect sizes for the functional outcome measures, the former position was adopted (ie, decline rates would follow established patterns). However, for the crucial everyday measures of IADL and ADL performance, the observed decline rates were significantly below established population norms. At 12 months, only 25% of participants experienced a 2-point or greater drop in the 36-point IADL scale, while by 24 months 28% had experienced this small increase in dependency. For the 30-point ADL performance scale, 6% were more dependent at 12 months and 8% at 24 months. Prior longitudinal research on cognitively demanding measures of everyday functioning indicates that age-related decline occurs later for these tasks than for the more basic abilities that were the focus of training. Reliable age-related decline on everyday problem-solving tasks has been shown not to occur until individuals are in their mid seventies, whereas declines on basic abilities such as reasoning and memory typically occur in their mid sixties.58

In summary, it is clear that proximal training effects occurred, that they continued (albeit at lower levels) through 24 months, and that a significant segment of trained individuals went forward through 2 years of life with better cognitive skills than did the controls. Due to lack of functional declines thus far, it is not yet clear whether differential functional decline across treatment groups will be observed in the future as this select cohort enters more fully into an age of functional loss.

Author Affiliations: Department of Psychology, University of Alabama at Birmingham (Dr Ball); National Institute on Aging (Dr Berch), National Institute of Nursing Research (Dr Helmers and Leveck), and National Heart, Lung, and Blood Institute (Dr Jobe), National Institutes of Health, Bethesda, Md; Institute on Aging and Departments of Health Policy and Epidemiology and Clinical and Health Psychology, University of Florida, Gainesville (Dr Marsiske); Hebrew Rehabilitation Center for the Aged, Roslindale, Mass (Dr Morriss); Department of Mental Hygiene, Johns Hopkins University, Baltimore, Md (Dr Rebok); Department of Medicine (Dr Smith) and Department of Psychiatry (Dr Unverzagt), Indiana University School of Medicine, Bloomington; New England Research Institutes, Watertown, Mass (Dr Tennstedt); and Department of Human Development and Family Studies, Pennsylvania State University, University Park (Dr Willis). Order of authorship on this article is alphabetical by author by vote of the ACTIVE Steering Committee.

Financial Disclosure: Dr Ball owns stock in Visual Awareness Inc, which owns the patent for Useful Field of View testing and training software.

Author Contributions: The New England Research Institutes served as the coordinating center for the study and handled all data analyses as determined by the consensus of the authors; each author of the study was principal investigator of his or her site, and all had access to the data. Study concept and design: Ball, Jobe, Leveck, Marsiske, Morriss, Rebok, Smith, Tennstedt, Unverzagt, Willis. Acquisition of data: Ball, Marsiske, Morriss, Rebok, Smith, Unverzagt, Willis.

Table 5. Training Effects on Proximal Outcomes Relative to Expected Annualized Declines*

<table>
<thead>
<tr>
<th>Source</th>
<th>Age Range, y</th>
<th>Magnitude of Change</th>
<th>Annualized Change, SD/y</th>
<th>ACTIVE Immediate Training Effect, SD</th>
<th>ACTIVE Training Effect at 2 y, SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schaie,53 1996</td>
<td>67-81</td>
<td>0.25 SD over 7 y (semantic lists)</td>
<td>0.036</td>
<td>0.25</td>
<td>0.17</td>
</tr>
<tr>
<td>Small et al,²⁴ 1999</td>
<td>≥78</td>
<td>0.25 SD over 6 y (immediate word recall)</td>
<td>0.041</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zelinski and Burnright,²⁰ 1997</td>
<td>67-81</td>
<td>0.50 SD over 16 y (immediate word recall)</td>
<td>0.031</td>
<td>0.031</td>
<td></td>
</tr>
<tr>
<td>Schaie,²⁰ 1996</td>
<td>67-81</td>
<td>0.22 SD over 7 y (semantic lists)</td>
<td>0.036</td>
<td>0.25</td>
<td>0.17</td>
</tr>
<tr>
<td>Ball and Owsley,⁸ 2000</td>
<td>≥55</td>
<td>0.16 SD over 2 y</td>
<td>0.08</td>
<td>1.46</td>
<td>0.87</td>
</tr>
</tbody>
</table>

*ACTIVE Indicates Advanced Cognitive Training for Independent and Vital Elderly.
COGNITIVE TRAINING INTERVENTIONS WITH OLDER ADULTS


47. Dudek FJ. The continuing misinterpretation of the standard error of measurement. Psychol Bull. 1979; 86:335-337.
52. Edwards JD, Wadley VG, Myers RS, Roenker DL, Cissell GM, Ball KK. The transfer of a speed of processing intervention to near and far cognitive functions. Gerontology. In press.

This we know: The earth does not belong to man, man belongs to the Earth. All things are connected like the blood that unites us all: Man did not weave the web of life, he is but a strand in it. Whatever he does to the web he does to himself.
—Chief Seattle (17867-1866)