The Mental Activity and eXercise (MAX) Trial

**A Randomized Controlled Trial to Enhance Cognitive Function in Older Adults**

Deborah E. Barnes, PhD, MPH; Wendy Santos-Modesitt, MA; Gina Poelke, PhD; Arthur F. Kramer, PhD; Cynthia Castro, PhD; Laura E. Middleton, PhD; Kristine Yaffe, MD

**Importance:** The prevalence of cognitive impairment and dementia are projected to rise dramatically during the next 40 years, and strategies for maintaining cognitive function with age are critically needed. Physical or mental activity alone result in relatively small, domain-specific improvements in cognitive function in older adults; combined interventions may have more global effects.

**Objective:** To examine the combined effects of physical plus mental activity on cognitive function in older adults.

**Design:** Randomized controlled trial with a factorial design.

**Setting:** San Francisco, California.

**Participants:** A total of 126 inactive, community-residing older adults with cognitive complaints.

**Interventions:** All participants engaged in home-based mental activity (1 h/d, 3 d/wk) plus class-based physical activity (1 h/d, 3 d/wk) for 12 weeks and were randomized to either mental activity intervention (MA-I; intensive computer) or mental activity control (MA-C; educational DVDs) plus exercise intervention (EX-I; aerobic) or exercise control (EX-C; stretching and toning); a 2 x 2 factorial design was used so that there were 4 groups: MA-I/EX-I, MA-I/EX-C, MA-C/EX-I, and MA-C/EX-C.

**Main Outcome Measures:** Global cognitive change based on a comprehensive neuropsychological test battery.

**Results:** Participants had a mean age of 73.4 years; 62.7% were women, and 34.9% were Hispanic or nonwhite. There were no significant differences between the groups at baseline. Global cognitive scores improved significantly over time (mean, 0.16 SD; P < .001) but did not differ between groups in the comparison between MA-I and MA-C (ignoring exercise, P = .17), the comparison between EX-I and EX-C (ignoring mental activity, P = .74), or across all 4 randomization groups (P = .26).

**Conclusions and Relevance:** In inactive older adults with cognitive complaints, 12 weeks of physical plus mental activity was associated with significant improvements in global cognitive function with no evidence of difference between intervention and active control groups. These findings may reflect practice effects or may suggest that the amount of activity is more important than the type in this subject population.

**Trial Registration:** clinicaltrials.gov Identifier: NCT00522899

**See Invited Commentary at end of article**

**During the next 40 years, it is anticipated that there will be an epidemic of dementia worldwide, with a 3- to 4-fold increase in the number of prevalent cases owing to longer life expectancies and demographic changes.**

Current treatments provide some symptomatic relief but do not alter the course of the disease, and several therapies that initially appeared promising have recently failed in phase III clinical trials. Attention is turning toward identification of preclinical disease and development of treatments to prevent or delay the onset of dementia; however, there is also concern about potential adverse effects of pharmacological treatment in individuals who may never become symptomatic. Behavioral interventions offer a potential strategy to prevent or delay dementia onset with minimal adverse effects in asymptomatic individuals.

There is growing evidence that both physical and mental activity can improve cognitive function in the short term and may lower the risk of developing dementia over the long term. Numerous observational studies and several systematic reviews have found that older adults who engage in mental or physical activity are less likely to experience cog-
nitive decline or develop dementia, and several have suggested that physical and mental activity may have independent or additive effects.\textsuperscript{7,11} However, randomized controlled trials (RCTs) are needed to establish a clear causal relationship between engaging in physical or mental activity and cognitive benefits.

To date, RCTs of physical and mental activity interventions have had mixed results.\textsuperscript{21} Mental activity interventions in both healthy older adults\textsuperscript{22,23} and individuals with mild cognitive impairment\textsuperscript{24,25} have typically found domain-specific improvements, in which benefits are observed in the specific cognitive activities trained with little evidence of generalization to other activities or domains. Exercise interventions in both healthy older adults\textsuperscript{26,27} and individuals with mild cognitive impairment\textsuperscript{28-30} have found that aerobic exercise and resistance training are associated with small to moderate improvements in cognitive function, particularly measures of attention, processing speed and executive function.\textsuperscript{31-33}

Taken together, these findings suggest that behavioral interventions that combine physical and mental activity may have more global effects than either physical or mental activity alone. However, few RCTs have studied the effects of physical and mental activity together.\textsuperscript{34} Therefore, we performed an RCT with a 2 × 2 factorial design to compare the effects of different physical and mental activity combinations on cognitive function in community-residing older adults with self-reported cognitive complaints.

METHODS

PARTICIPANTS

Study participants were recruited primarily through direct mailing to the neighborhoods adjacent to the intervention site (Stonestown YMCA, San Francisco, California), as well as advertisements, fliers, physician and friend referrals, and recruitment databases. Inclusion criteria were age 65 years or older, cognitive complaint (defined as answering yes to the question “Do you feel that your memory or thinking skills have gotten worse recently?”), English language fluency, not currently engaging in aerobic exercise or intensive computer training (≤2 d/wk, ≤30 minutes per session in the past 3 months), and not planning to travel for more than 1 week during the study period. Exclusion criteria were dementia (based on self-report, physician diagnosis, or scoring 18 or less on the modified Telephone Interview for Cognitive Status\textsuperscript{35}), other neurological or major psychiatric disorder, significant heart or lung disease, limited life expectancy, or other factors that could potentially limit ability to participate fully in the intervention.

All study procedures were approved by the Committee on Human Research at the University of California, San Francisco, and the Research and Development Committee at the San Francisco Veterans Affairs Medical Center. All participants provided written informed consent and received physician approval to participate.

INTERVENTIONS

Study participants were randomized to both a home-based mental activity intervention (MA-I) or mental activity control (MA-C) group and a class-based exercise intervention (EX-I) or exercise control (EX-C) group in a 2 × 2 factorial design. Active control groups were used to account for the effects of factors such as interaction with a computer and social interaction during a group exercise class.

Mental Activity

All participants were provided with detailed written and in-person verbal instructions regarding their assigned mental activities, which were performed independently at home on a computer for 60 min/d, 3 d/wk for 12 weeks. The MA-I group performed games designed to enhance the speed and accuracy of visual and auditory processing (Posit Science). For the first 6 weeks, games focused on visual tasks, including determining the direction of visual sweeps, identifying bird pairs, tracking the location of moving gems, and identifying targets in peripheral vision. For the second 6 weeks, games focused on auditory tasks, including determining the direction of auditory sweeps, distinguishing between similar sounds, matching sound pairs, recalling a series of sounds, following verbal instructions, and answering questions about verbal stories. Program difficulty adjusted automatically and continuously based on each participant’s level of performance, to maintain a correct response level of approximately 85%.

The MA-C group watched DVDs of educational lectures on art, history, and science. After each session, participants answered approximately 6 paper-based, multiple-choice or short-answer lecture-specific questions. Participants watched the DVDs on a computer for 60 min/d, 3 d/wk for 12 weeks to match the conditions of the MA-I group.

Exercise

All participants attended study-specific group exercise classes at a local YMCA for 60 min/d, 3 d/wk for 12 weeks. The EX-I class consisted of 10 minutes of warm-up, 30 minutes of aerobic exercise (standard dance-based aerobics format), 5 minutes of cooldown, 10 minutes of strength training, and 5 minutes of stretching and relaxation. Heart rates were monitored by having participants check their wrist or neck pulse for 10 seconds at the beginning, peak, and end of class and record the values in an exercise journal, with a target peak heart rate of 60% to 75% of the maximum for the participant’s age.

The EX-C class consisted of 10 minutes of warm-up, 30 minutes of stretching and toning, 10 minutes of strength training, and 10 minutes of relaxation. Heart rate was monitored as described for the EX-I class, but with a goal of not raising heart rates above resting levels. All classes were taught by a single certified exercise instructor with experience conducting classes for elderly adults and with a maximum of 12 class participants at any time.

Compliance and adverse events were monitored using weekly journals and biweekly telephone check-ins, and motivational counseling was provided if compliance fell below 80%. Daily attendance was also recorded for the exercise classes.

RANDOMIZATION AND BLINDING

Participants were randomized in blocks of 4. The randomization sequence was prepared in advance by using a random-number generator on a computer. Research staff involved with enrollment and outcome assessment were unaware of the randomization sequence and blinded to group assignment. Study participants were unaware of study hypotheses and were told that the goal of the study was to compare the effects of different physical and mental activity programs.
OUTCOMES

Primary Outcome

The primary outcome of the study was 12-week change in cognitive function based on a composite score from a comprehensive neuropsychological test battery that included 6 tests. Specific tests were selected because they have good validity and reliability and are sensitive to cognitive decline in older adults and included measures of verbal learning and memory (Rey Auditory Verbal Learning Test [RAVLT]), verbal fluency (letter and category), processing speed (Digit Symbol Substitution Test [DSST]), executive function/mental flexibility (Trail-Making Test, parts A and B [Trails A and B]), executive function/inhibition (Eriksen Flanker Test [EFT], congruent and incongruent reaction times), visuospatial function (Useful Field of View [UFOV], processing speed, divided attention, and selective attention).

The Modified Mini-Mental State examination was performed only at baseline to assess global cognitive status (range, 0-100).

A composite cognitive score was created by converting all individual cognitive scores to standardized z scores by subtracting the baseline group mean and dividing by the baseline group SD, and then averaging the standardized z scores. For tests with more than 1 component, the component scores were first averaged to create a single z score for each test (eg, z scores for Trails A and B were averaged to create a single Trails z score) so that all tests were weighted equally, and then the mean z scores for the 6 tests was calculated. In participants who were missing data for no more than 3 cognitive tests, the composite score was based on nonmissing tests.

Other Measures

Demographic measures included age, sex, race, ethnicity, and annual income. Comorbid medical conditions were determined based on self-report of prior physician diagnosis. Physical performance was assessed with the Senior Fitness Test, which includes the following standard measures: chair stand, arm curl, 2-minute step, sit and reach, back scratch, and 8-foot up-and-go tests.

POWER

Our study had 80% power to detect an effect size of approximately 0.3 SD when examining the main effects of MA-I vs MA-C (ignoring exercise) and EX-I vs EX-C (ignoring mental activity) (63 participants per group) and an effect size of 0.45 SD for comparisons across the 4 randomization groups (31 participants per group), assuming 80% correlation between measures within participants (2-sided α = .05).

STATISTICAL ANALYSIS

All analyses were performed using intent-to-treat principles. Baseline characteristics were compared between groups with the use of χ2 tests for categorical variables and analysis of variance for continuous variables. Mixed effects linear regression models were used to examine change in cognitive scores as a function of randomization group, time (baseline or postintervention), and the group × time interaction. Random intercepts for each participant were used to model correlation between participants over time. Robust estimation of the variance-covariance structure was used to account for the possibility of unequal variance at the 2 time points. Because of the factorial design, analyses were performed to compare MA-I vs MA-C (ignoring exercise) and EX-I vs EX-C (ignoring mental activity), as well as all 4 randomization groups. Analyses were performed using the xtreg command in Stata 12 software (StataCorp).

Secondary analyses examined the effects of the interventions on individual cognitive tests. Based on prior studies, we hypothesized a priori that the effects of EX-I would be greatest for measures of executive function, especially the EFT incongruent task, and that the effects of the MA-I would be greatest for the UFOV test, which is similar to one of the games in the training program.

To help interpret the main results of the study, we also performed 2 post hoc subgroup analyses to assess whether associations differed as a function of low memory scores (defined as delayed recall ≤5 words [approximately 1 SD below the mean] on the RAVLT) or poor physical function (defined as <65 full steps on the 2-minute step test) at baseline, by including 3-way interaction terms for these variables, randomization group, and time in the random effects regression models.

All results were similar when analyses were restricted to participants who completed the study; therefore, only intent-to-treat results are shown.
MA-C groups, the EX-I and EX-C groups, or all 4 randomization groups.

Baseline cognitive test scores by randomization group are shown in Table 2. On average, study participants had moderate to high levels of cognitive function at baseline, consistent with their age and education levels. There were no significant differences in baseline cognitive test scores between the MA-I and MA-C groups, the EX-I and EX-C groups, or all 4 randomization groups.

The effects of the interventions on our primary outcome of cognitive change in the composite score are shown in Figure 2. There was a significant main effect of time, indicating significant improvement in all of the groups (mean, 0.16 SD; P < .001). However, there were no significant differences when we compared MA-I vs MA-C (P = .17), EX-I vs EX-C (P = .74), or all 4 randomization groups (P = .26).

When we examined individual cognitive tests, there were significant main effects for time for the Digit Symbol Substitution Test, Trails A, EFT congruent and incongruent reaction times, and UFOV divided and selective attention, with trends toward improvement on most other tests, except delayed recall on the RAVLT (Figure 3). The only significant differences between the groups were observed for the...
Approximately 27% of participants (n = 34) had low memory scores at baseline. The 3-way interaction between baseline memory, mental activity group, and time (data not shown).

In this 2 × 2 factorial RCT examining the effects of different physical and mental activity interventions on cognitive function in nondemented, inactive older adults with cognitive complaints, we found that cognitive scores improved significantly over the course of 12 weeks, but there were no significant differences between the intervention and active control groups. These results may suggest that, in this study population, the amount of activity is more important than the type of activity, because all groups participated in both mental activity and exercise for 60 min/d, 3 d/wk for 12 weeks. Alternatively, the cognitive improvements observed may be due to practice effects.

To assess the possibility that our results were due to practice, we recruited an additional 12 participants to participate in a no-contact control study in which all study procedures were identical except that there was no intervention. Composite cognitive test scores in these participants improved 0.08 SD compared with 0.16 SD in the main study, suggesting that some, but not all, of the improvements observed may have been due to repeated testing.

Our findings differ from those of prior RCTs, which have found that a similar intensive computer training program improved cognitive function more than educational DVDs in healthy older adults. This may be attributable to differences in the outcome measures used or the manner in which the intervention was implemented. Consistent with prior studies, we found that the MA-I training yielded significantly greater improvements in those cognitive tests that were similar to the training program.

Abbreviations: DSST, Digit Symbol Substitution Test; EFT, Eriksen Flanker Test; RAVLT, Rey Auditory Verbal Learning Test; Trails A and B, Trail-Making Test, parts A and B; UFOV, useful field of view.

There were no significant differences between the mental activity intervention and mental activity control groups (ignoring exercise), the exercise intervention and exercise control groups (ignoring mental activity), or all 4 randomization groups.

UFOV divided and selective attention tasks, which improved more in the MA-I than in the MA-C group (Table 3).

Approximately 27% of participants (n = 34) had low memory scores at baseline. The 3-way interaction between baseline memory, mental activity group, and time approached but did not reach statistical significance (P = .054) and, in participants with low memory scores, the difference in cognitive change between the MA-I and MA-C groups also approached statistical significance (P = .08) (Figure 4). There was no evidence of interaction between baseline physical function, randomization group, and time (data not shown).

In this 2 × 2 factorial RCT examining the effects of different physical and mental activity interventions on cognitive function in nondemented, inactive older adults with cognitive complaints, we found that cognitive scores improved significantly over the course of 12 weeks, but there were no significant differences between the intervention and active control groups. These results may suggest that, in this study population, the amount of activity is more important than the type of activity, because all groups participated in both mental activity and exercise for 60 min/d, 3 d/wk for 12 weeks. Alternatively, the cognitive improvements observed may be due to practice effects.

To assess the possibility that our results were due to practice, we recruited an additional 12 participants to participate in a no-contact control study in which all study procedures were identical except that there was no intervention. Composite cognitive test scores in these participants improved 0.08 SD compared with 0.16 SD in the main study, suggesting that some, but not all, of the improvements observed may have been due to repeated testing.

Our findings differ from those of prior RCTs, which have found that a similar intensive computer training program improved cognitive function more than educational DVDs in healthy older adults. This may be attributable to differences in the outcome measures used or the manner in which the intervention was implemented. Consistent with prior studies, we found that the MA-I training yielded significantly greater improvements in those cognitive tests that were similar to the training program.
Learning Test (RAVLT), total words learned; the Trail-Making Test, part B (Trails B); and UFOV processing speed (UFOV-PS). DSST; Trail-Making Test, part A (Trails A); Eriksen Flanker Test (EFT), congruent and incongruent reaction times; Useful Field of View divided attention parts A and B; UFOV, Useful Field of View.

Figure 3. Change in individual cognitive tests. Cognitive function improved significantly (\(* P < .05\)) for the following scores: Digit Symbol Substitution Test (DSST), Trail-Making Test, part A (Trails A); Eriksen Flanker Test (EFT), congruent and incongruent reaction times; Useful Field of View divided attention (UFOV-DA) and selective attention (UFOV-SA); and composite score. Improvements approached statistical significance (\(* P < .10\)) for Rey Auditory Verbal Learning Test (RAVLT), total words learned; the Trail-Making Test, part B (Trails B); and UFOV processing speed (UFOV-PS).

Table 3. Standardized Mean Change in Cognitive Scores by Randomization Group

<table>
<thead>
<tr>
<th>Cognitive Test</th>
<th>Exercise Control (n = 32)</th>
<th>Exercise Intervention (n = 31)</th>
<th>Exercise Control (n = 31)</th>
<th>Exercise Intervention (n = 32)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.26</td>
</tr>
<tr>
<td>Verbal learning and memory (RAVLT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of words learned</td>
<td>0.33 (0.09 to 0.58)</td>
<td>0.14 (−0.14 to 0.43)</td>
<td>0.13 (−0.11 to 0.37)</td>
<td>−0.04 (−0.42 to 0.33)</td>
<td>.38</td>
</tr>
<tr>
<td>No. of words recalled</td>
<td>0.01 (−0.32 to 0.33)</td>
<td>0.02 (−0.28 to 0.32)</td>
<td>−0.10 (−0.36 to 0.16)</td>
<td>−0.07 (−0.46 to 0.32)</td>
<td>.93</td>
</tr>
<tr>
<td>Verbal fluency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of words, by letter</td>
<td>−0.05 (−0.33 to 0.24)</td>
<td>0.08 (−0.21 to 0.37)</td>
<td>0.24 (−0.11 to 0.58)</td>
<td>0.22 (−0.15 to 0.58)</td>
<td>.57</td>
</tr>
<tr>
<td>No. of words, by category</td>
<td>0.06 (−0.26 to 0.38)</td>
<td>−0.07 (−0.41 to 0.26)</td>
<td>0.22 (−0.06 to 0.50)</td>
<td>0.18 (−0.24 to 0.66)</td>
<td>.59</td>
</tr>
<tr>
<td>Processing speed (DSST)</td>
<td>0.15 (−0.02 to 0.32)</td>
<td>0.19 (0.04 to 0.33)</td>
<td>0.27 (0.03 to 0.51)</td>
<td>0.08 (−0.13 to 0.30)</td>
<td>.71</td>
</tr>
<tr>
<td>Executive function/mental flexibility, s(^c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trails A</td>
<td>−0.36 (−0.58 to −0.15)</td>
<td>−0.12 (−0.32 to 0.07)</td>
<td>−0.03 (−0.50 to 0.44)</td>
<td>−0.36 (−0.63 to −0.08)</td>
<td>.24</td>
</tr>
<tr>
<td>Trails B</td>
<td>−0.22 (0.45 to 0.002)</td>
<td>−0.18 (−0.49 to 0.13)</td>
<td>0.13 (−0.21 to 0.48)</td>
<td>−0.25 (−0.51 to 0.01)</td>
<td>.31</td>
</tr>
<tr>
<td>Executive function/inhibition (EFT), ms(^c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congruent reaction time</td>
<td>−0.17 (−0.51 to 0.16)</td>
<td>0.01 (−0.42 to 0.43)</td>
<td>−0.17 (−0.41 to 0.06)</td>
<td>−0.33 (−0.55 to −0.11)</td>
<td>.51</td>
</tr>
<tr>
<td>Incongruent reaction time</td>
<td>−0.12 (−0.36 to 0.12)</td>
<td>−0.07 (−0.48 to 0.33)</td>
<td>−0.15 (−0.52 to 0.22)</td>
<td>−0.33 (−0.58 to −0.08)</td>
<td>.60</td>
</tr>
<tr>
<td>Visuospatial function (UFOV), ms(^c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Processing speed</td>
<td>−0.23 (−0.66 to 0.19)</td>
<td>−0.04 (−0.43 to 0.34)</td>
<td>−0.17 (−0.61 to 0.26)</td>
<td>−0.39 (−0.77 to −0.02)</td>
<td>.65</td>
</tr>
<tr>
<td>Divided attention</td>
<td>−0.13 (−0.48 to 0.22)</td>
<td>−0.17 (−0.39 to 0.05)</td>
<td>−0.60 (−0.99 to −0.22)</td>
<td>−0.62 (−0.97 to −0.26)</td>
<td>.05</td>
</tr>
<tr>
<td>Selective attention</td>
<td>−0.18 (−0.59 to 0.22)</td>
<td>−0.14 (−0.43 to 0.16)</td>
<td>−0.34 (−0.58 to −0.11)</td>
<td>−0.71 (−0.96 to −0.46)</td>
<td>.02</td>
</tr>
</tbody>
</table>

Abbreviations: DSST, Digit Symbol Substitution Test; EFT, Eriksen Flanker Test; RAVLT, Rey Auditory Verbal Learning Test; Trails A and B, Trail-Making Test, parts A and B; UFOV, Useful Field of View.

\(^a\) Calculated from mixed effects linear regression models that included terms for randomization group, time (baseline or postintervention) and the group \(\times\) time interaction, \(P\) values reflect the statistical significance of the group \(\times\) time interaction across the 4 randomization groups.

\(^b\) There were significant differences between mental activity intervention and mental activity control groups (ignoring exercise) for UFOV, divided attention and selective attention. There were no other significant differences between the mental activity intervention and mental activity control groups (ignoring exercise), the exercise intervention and exercise control groups (ignoring mental activity), and all 4 randomization groups.

\(^c\) For timed tests, negative changes reflect improvement (ie, faster performance).
There is growing evidence that even these less intensive interventions may have cognitive and physical benefits in older adults. Future studies should consider inclusion of both active and no-control conditions.

We did not observe any differences between the EX-I (aerobic) and EX-C (stretching and toning) groups for either global cognitive function or individual cognitive tests. This differs from prior RCTs in healthy older adults, which have found that aerobic exercise improves cognitive function and increases hippocampal volume compared with stretching/toning. It is possible that our 12-week intervention was not long enough or intense enough to achieve a substantially greater aerobic response in the intervention group and that a difference between the groups would have emerged in a longer study. It is also possible that participants in the stretching/toning group engaged in more aerobic activities outside the study as a function of their participation and readiness to exercise. Alternatively, it is possible that in our study population of older adults with cognitive complaints, both aerobic exercise and stretching/toning resulted in physical changes that were equally beneficial. This hypothesis is supported by findings in other recent RCTs indicating that resistance training alone is associated with improvements in executive function.

STRENGTHS AND LIMITATIONS

Our study had several key strengths, including the 2 × 2 factorial design and the use of active control conditions, which enabled us to control for factors such as social interaction during the group exercise class and mental stimulation associated with using a computer. However, there is growing evidence that even these less intensive interventions may have cognitive and physical benefits in older adults. Future studies should consider inclusion of both active and no-contact control conditions.

Our study also has several potential limitations. Although more than one-third of participants were Hispanic or nonwhite, most were highly educated, raising concern about the generalizability of the findings. In addition, our study did not include objective measures of aerobic fitness such as peak oxygen consumption; therefore, we cannot be sure that participants in the EX-I group achieved improvements in aerobic fitness. Our cognitive test battery also did not include the digit span test, which was found in a recent review to be most consistently affected by exercise. Finally, we did not perform full clinical evaluations of study participants.

In summary, in inactive older adults with cognitive complaints, 12 weeks of a combined mental activity and exercise program was associated with significant improvements in global cognitive function, with no evidence of differences between intervention and active control groups. These findings may reflect practice effects or may suggest that the amount of activity is more important than the type in this subject population.