



Published in final edited form as:

Ann Behav Med. 2008 December ; 36(3): 280–291. doi:10.1007/s12160-008-9064-5.

Exercise, Fitness, and Neurocognitive Function in Older Adults: The “Selective Improvement” and “Cardiovascular Fitness” Hypotheses

Ann L. Smiley-Oyen, Ph. D.,

Department of Kinesiology, Iowa State University, 235 Forker Building, Ames, IA 50011, USA
asmiley@iastate.edu

Kristin A. Lowry, M.S, P.T.,

Department of Kinesiology, Iowa State University, 235 Forker Building, Ames, IA 50011, USA

Sara J. Francois, D.P.T.,

Department of Physical Therapy, Mary Greeley Medical Center, 1111 Duff Avenue, Ames, IA 50010, USA

Marian L. Kohut, Ph. D., and

Department of Kinesiology, Iowa State University, 235 Forker Building, Ames, IA 50011, USA

Pantelimon Ekkekakis, Ph. D.

Department of Kinesiology, Iowa State University, 235 Forker Building, Ames, IA 50011, USA

Abstract

Background—Although basic research has uncovered biological mechanisms by which exercise could maintain and enhance adult brain health, experimental human studies with older adults have produced equivocal results.

Purpose—This randomized clinical trial aimed to investigate the hypotheses that (a) the effects of exercise training on the performance of neurocognitive tasks in older adults is selective, influencing mainly tasks with a substantial executive control component and (b) performance in neurocognitive tasks is related to cardiorespiratory fitness.

Methods—Fifty-seven older adults (65–79 years) participated in aerobic or strength-and-flexibility exercise training for 10 months. Neurocognitive tasks were selected to reflect a range from little (e.g., simple reaction time) to substantial (i.e., Stroop Word–Color conflict) executive control.

Results—Performance in tasks requiring little executive control was unaffected by participating in aerobic exercise. Improvements in Stroop Word–Color task performance were found only for the aerobic exercise group. Changes in aerobic fitness were unrelated to changes in neurocognitive function.

Conclusions—Aerobic exercise in older adults can have a beneficial effect on the performance of speeded tasks that rely heavily on executive control. Improvements in aerobic fitness do not appear to be a prerequisite for this beneficial effect.

Keywords

Stroop; Wisconsin Card Sort Test; Cardiorespiratory fitness; Executive processing; Aerobic exercise

Introduction

Aging is associated with considerable declines in a wide spectrum of cognitive abilities, including speed and accuracy of perception, decision-making, task-switching, working memory, and multitasking [1,2]. These declines, particularly when they escalate to dementia, can have a dramatic impact on the independence, safety, activities of daily living, and overall quality of life of older adults.

Several prospective epidemiologic [3–8], cross-sectional [9,10], and experimental [11,12] studies published in recent years have brought to the forefront the notion that physical activity and exercise might be effective in slowing the rate, or even reversing, the cognitive decline associated with aging. Building upon earlier work that had established a positive effect of exercise on brain vascularization [13–15], recent animal studies have also demonstrated a role in neurogenesis [16,17] and the upregulation of neurotrophic factors [18–21]. A recent study of 11 adult human participants also showed that blood volume in the dentate gyrus of the hippocampus (the only hippocampal subregion that supports adult neurogenesis), assessed by magnetic resonance imaging as an *in vivo* marker of neurogenesis, increased significantly over a 3-month period of aerobic exercise. Furthermore, this increase in dentate gyrus blood volume was significantly correlated with gains in maximal aerobic capacity, as well as with the improvement in short-term memory in an auditory verbal learning test [22]. By highlighting several plausible biological mechanisms, this body of evidence has laid the foundation for establishing a causal link between exercise and the preservation or improvement of neurocognitive function in aging. However, certain important issues remain unresolved.

Experimental studies, in which participants were randomly assigned to conditions, have yielded equivocal results, with some showing that aerobic-type activities (e.g., walking) resulted in significant benefits compared to controls (e.g., [12,23–26]) and others producing null findings (e.g., [27–31]). In some cases, the null findings were attributed to such methodological factors as the already-high level of cognitive functioning of the participants at baseline or the short intervention periods that might have been inadequate for substantial changes in fitness to occur (e.g., [27,32]).

Kramer and collaborators [12] proposed a hypothesis that offers a new and intriguing perspective on the conflicting findings. Based on the frontal lobe hypothesis of cognitive aging [33,34], which suggests that the loss of brain tissue and corresponding declines in neurocognitive function are not uniform across brain regions, Kramer proposed that exercise, particularly aerobic exercise that results in measurable gains in cardiorespiratory fitness, leads to selective, rather than generalized, benefits. Specifically, tasks with a substantial frontal-lobe-dependent executive control component exhibit the largest aging-related declines and should also be expected to show the largest exercise- or fitness-related improvements. Such tasks include those that involve use of information retained in working memory, simultaneous execution of multiple tasks, task-switching, and inhibition of an ongoing or prepotent response. On the other hand, tasks that do not require a substantial executive control component, such as simple reaction time, are less affected by aging and are, thus, also less likely to benefit from exercise or fitness.

Evidence in support of this “selective improvement” hypothesis has come from a training study [12] and a subsequent meta-analysis of randomized trials [35]. In the 6-month training study [12], 124 older adults (60–75 years of age) were randomly divided into an aerobic (walking) and a non-aerobic (stretching and toning) group. The former showed modest but significant gains in aerobic fitness (+5.1%), whereas the latter did not (–2.8%). Before and after these treatments, the participants underwent testing on neurocognitive tasks specifically selected to

rely heavily (e.g., task-switching, responding to “incompatible” stimuli, or stopping an ongoing action) or not-so-heavily (e.g., simple reaction time, involving “compatible” stimuli, without switching or stopping) on executive control. Consistent with the “selective improvement” hypothesis, the results showed that, although the two groups did not differ in the low-executive-control tasks, the aerobic group showed improvements in the high-executive-control tasks. The subsequent meta-analysis examined 18 intervention trials that involved adults from 55–80 years of age and supervised aerobic training with random assignment to conditions [35]. The meta-analysis examined four competing theoretical predictions, namely that the effects of aerobic exercise or fitness would be specific to tasks characterized by (a) speed (tasks representing low-level neurological functioning, such as simple reaction time), (b) visuospatial ability (ability to transform or remember visual and spatial information, such as redrawing shapes from memory), (c) controlled processes (tasks requiring some cognitive control, such as simple rule-based decision-making or choice reaction time tasks), and (d) executive control (planning, inhibition, and scheduling, such as responding to one cue while suppressing other, simultaneously presented, conflicting, or irrelevant cues). Exercise training groups showed an average effect size of 0.48, whereas control groups showed an average effect size of 0.16. Again providing support for the “selective improvement” hypothesis, the effect was found to be the strongest for executive-control tasks (0.68), followed by controlled processes (0.46), visuospatial tasks (0.43), and speed tasks (0.27).

Kramer and coworkers have maintained that aerobic exercise and associated gains in aerobic or cardiovascular fitness, of at least small magnitude, are necessary for the selective benefits to occur [36–38]. However, the evidence for this claim is inconclusive. In fact, the aforementioned meta-analysis [35] found no relationship between the magnitude of gains in aerobic fitness (coded as unreported; moderate, 5–11%; and large, 12–25%) and the effect of exercise interventions on neurocognitive function. Similarly, a more extensive meta-regression analysis that focused specifically on this “cardiovascular fitness” hypothesis showed (a) no relationship between fitness and cognitive performance among older adults in cross-sectional studies, (b) no relationship between fitness and cognitive performance at the end of intervention trials, and (c) a negative relationship between gains in fitness and changes in cognitive performance from before to after intervention trials [39]. Age interacted with fitness, in that fitness was a significant negative predictor of cognitive performance specifically among older adults (and was unrelated to cognitive performance among children, young adults, and middle-aged adults). The present study extends this line of research. The primary purpose was to test the “selective improvement” hypothesis. The neurocognitive tests that were administered were specifically selected to reflect different degrees of frontally mediated executive control. The second purpose of the present study was to examine the “cardiovascular fitness” hypothesis by investigating whether changes in performance in the neurocognitive tasks were related to changes in aerobic fitness. We employed similar interventions as those in the Kramer et al. [12] study (aerobic exercise versus stretching-and-toning) but of longer duration (10 instead of 6 months). The longer duration was intended to increase the power of the intervention by (a) providing larger amounts of exercise, potentially leading to a larger separation between the distributions of the two groups in relevant dependent variables and (b) producing highly variable changes in aerobic fitness, conducive to detecting any covariation with changes in neurocognitive function.

Method

Participants

The present study was part of a larger randomized clinical trial (RCT) on the effects of exercise on immune function [40]. Institutional Review Boards at participating academic and medical institutions approved all experimental procedures. To participate in the RCT, volunteers had

to be 64 years of age or older and living independently. Prospective participants were excluded on the basis of health-related reasons and for being too physically active and/or physically fit at baseline. The following health-related exclusionary criteria were used: (a) having an autoimmune disorder or other disease likely to impact the immune system, (b) having been diagnosed with cancer within the previous 5 years, (c) taking medication that affected the immune system (such as oral corticosteroids), (d) taking anxiolytic or antidepressant medications, or (e) suffering from conditions for which vigorous exercise is contraindicated, including uncontrolled metabolic or unstable cardiovascular disease (see Ref. [41], p. 50). The ability to perform exercise safely was also ascertained through a medically supervised graded treadmill test performed following randomization (see the “Cardiorespiratory Fitness” section below). Participants with electrocardiographic or other findings during the treadmill test were referred to their primary health care provider for additional testing and treatment. Furthermore, prospective participants were excluded if they were participating in exercise 3 or more times per week at 40% or more of their heart rate reserve, or if their aerobic fitness level was above the 75th percentile for their age and gender, as assessed by the 6-min walk component of the Senior Fitness Test [42,43].

Participation in the RCT was solicited by contacting older adults residing in two nearby communities in the Midwestern United States (combined three-county area population, according to the 2000 US Census, of 480,806, with 53,721 aged 65 years or older). The study was advertised through press releases in local newspapers, visits to senior citizen meal sites, churches and church groups, senior community service organizations, independent living housing facilities, and by friend referrals. Recruitment was done in two cohorts (first cohort of 50: August 2002 to October 2002; second cohort of 59: August 2003 to October 2003).

Of the 133 individuals who expressed interest and were screened for eligibility, 24 were excluded for not meeting the established criteria (see Fig. 1). Thus, 109 were subsequently randomized either to a cardiovascular exercise group (CARDIO, $n=55$) or to a group involved in flexibility exercises and weight-training (FLEX-TONE, $n=54$). Before group assignments were announced to the participants, they were determined by a research assistant who had no knowledge of or direct contact with the participants. Group allocation alternated between CARDIO and FLEX-TONE, while also balancing for gender, cardiorespiratory fitness, and the use of beta-blockers. Two individuals, one from each group, dropped out following randomization. After the preparticipation graded treadmill test (see below), 27 individuals were referred to their primary health care providers for additional testing and/or treatment (54% from CARDIO, 46% from FLEX-TONE). Of them, all but one received clearance to initiate participation in the RCT. Another individual continued to exhibit hypertensive symptoms and was excluded for precautionary reasons. No major adverse events or side effects associated with the treatments occurred.

Following the preparticipation graded treadmill tests but prior to the initiation of the treatments, the 105 remaining participants were asked if they wanted to enroll in the present study by taking part in a battery of neurocognitive tests. Of them, 78 individuals (74%) agreed to participate. One person was excluded for having a history of neurological disease (Parkinson's) and five individuals were excluded for being 80 years of age or older. Ten additional individuals were excluded over the course of the 10-month interventions: four of them dropped out of the interventions, three passed away, and three participants in the CARDIO group did not comply with the target heart rate prescription at least 50% of the time. “Intent to treat” analyses were performed with these three individuals included in the sample but none of the conclusions of the study were affected. Thus, to maintain the meaningfulness of comparisons between the treatment modalities (CARDIO versus FLEX-TONE), the results reported below refer to analyses without these three participants. Finally, five individuals were scheduled but did not return to complete all the required neurocognitive tests.

Thus, data for 57 participants (16 men, 41 women) were entered into the analyses (see Table 1). Of them, seven men and 21 women belonged to the CARDIO group and nine men and 20 women belonged to the FLEX-TONE group. For a summary of participant characteristics, see Table 1. All participants completed the written form of the Symbol Digit Modalities Test [44,45] at each testing session. They all scored within ± 1.5 SD of the normative mean for their age and years of education, indicating no cerebral impairment. Scores did not differ between participants in the CARDIO and FLEX-TONE groups.

Measures

Cardiorespiratory Fitness—Symptom-limited treadmill tests were performed before the initiation and after the completion of the interventions to assess cardiorespiratory fitness by a qualified clinical exercise physiologist and a supervising physician (both blinded to group assignments). A modified Bruce protocol was used. The standard Bruce protocol starts from 1.7 mph and 10% grade and proceeds in 3-min stages (2.5 mph, 12% grade; 3.4 mph, 14% grade; 4.2 mph, 16% grade, and so on). The modification used in the present study consisted of adding two stages at the start of the protocol (1.7 mph, 0% grade; 1.7 mph, 5% grade). The supervising physician monitored a continuous electrocardiogram. None of the tests was interrupted due to medical complications. Peak oxygen uptake (VO_2peak) was estimated from treadmill performance (time in minutes) using the general formula of Foster et al. [46], after subtracting 6 min for the two initial stages of the modified version¹:

$$\text{VO}_2\text{peak} \left(\text{ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1} \right) = 14.760 - 1.379(\text{time}) + 0.451(\text{time}^2) - 0.012(\text{time}^3)$$

All participants also underwent a 6-min walk test (as implemented in the Senior Fitness Test; [43,47]) both before the initiation and after the completion of the interventions. Estimated values of VO_2peak based on this test were used in the cases of three individuals who could not undergo postintervention graded treadmill tests due to scheduling conflicts. Changes in cardiorespiratory fitness for these individuals were determined on the basis of the VO_2peak estimates derived from the pre- and postintervention 6-min walk tests. The VO_2peak estimation formula suggested by the American College of Sports Medicine (Ref. [41], p. 70) was used for this purpose.

Neurocognitive Tests—The individuals administering the neurocognitive tests were blinded to group assignments. The selection of neurocognitive tests was driven by the “selective improvement hypothesis” outlined in the introduction [12]. Specifically, the different neurocognitive tests were conceptualized as lying on a continuum ranging from being minimally to being substantially dependent upon frontally mediated executive control. In doing so, we adhered closely to the rationale followed in Colcombe and Kramer's [35] meta-analysis, selecting tasks in a manner consistent with the four-group categorization utilized by these authors: (a) speeded, (b) visuospatial, (c) controlled, and (d) executive tasks.

Specifically, as a test representative of the *speeded* category (low-level neurological functioning), we selected a simple reaction time task. In the *visuospatial* category (transforming or remembering visual and spatial information), we included an incompatible 8-choice reaction time task, in which participants had to move to the button opposite of the lighted target. This task included an explicit speed component as well as an inhibition component. Tasks that fall

¹Data collected from a diverse sample of 230 individuals (including cardiac patients, healthy adults, and athletes) showed that the mean error (measured vs. predicted) based on this formula was $-0.6 \pm 3.4 \text{ ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$, the correlation between measured and predicted values was 0.97, and the standard error of prediction was $3.6 \text{ ml} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ [46].

under the *controlled* category require controlled attention but can be automated with sufficient practice. Consistent with Colcombe and Kramer's description, we included in this category several choice reaction time tasks (8-choice, Stroop Word, and Stroop Color). The final category, *executive*, includes processes such as inhibition, planning, and working memory, which require constant mediation by a central executor. The tasks chosen for this category were Go/No-Go and Stroop Word–Color conflict. Finally, we included a task that requires executive processing but does not include a speeded component, namely the Wisconsin Card Sort Test. We included this task in order to decipher whether the effects of exercise and/or fitness are larger on speeded tasks that require executive control or on tasks that are heavily dependent on executive processing but have no speed component. Each test is described in more detail below.

Procedures

Exercise Training—The exercise program took place at the same location for both the CARDIO and FLEX-TONE groups and was led by the same qualified exercise leader (obviously, blinding of the individual administering the treatments is not possible in the case of exercise). Both groups exercised three times per week for 10 months. Both the CARDIO and the FLEX-TONE programs started with a similar 10-min warm-up routine of stretching and mild aerobic exercises and ended with a 10-min cool down that involved stretching and some balance work.

After the warm-up, participants in the CARDIO group exercised for 25–30 min on the aerobic exercise equipment of their choice (including treadmills, stepper machines, arm ergometers, stationary cycles, vertical climbing machines, cross trainers, and elliptical machines). Heart rate during exercise was monitored continuously with heart rate monitors. The participants were given target heart rate prescriptions based on the peak heart rate achieved during their graded treadmill tests. Individualized prescriptions started at 45–60% of heart rate reserve, progressed to 60–70%, and were then maintained at 65–80% for the remainder of the 10-month intervention period ([41], p. 141). Participants on beta-blockers regulated exercise intensity mainly using the rating of perceived exertion (RPE; [41], p. 146). They advanced their effort in a manner consistent with the prescription identified above, eventually progressing to a “Somewhat hard” (i.e., rating of “13”) to “Hard” (i.e., rating of “15”) level. All CARDIO participants recorded the time they exercised, the machines they used, and the heart rate and RPE ratings achieved during each workout.

The FLEX-TONE group also met three times per week for 10 months. The participants in this group performed strength, flexibility, and balance exercises. After the warm-up, elements of yoga and Tai Chi, Flex bands, free hand weights, and stability balls were used for 25–30 min. Resistance training exercises using weight machines designed to train all major muscle groups were used during the second half of the intervention period. Consistent with American College of Sports Medicine guidelines ([41], pp. 154–158), eight to ten exercises and 10–15 repetitions per exercise were performed.

Neurocognitive Tests—The following types of neurocognitive tests were administered: (a) reaction-time tests, (b) the Stroop test, and (c) the Wisconsin Card Sort Test. The participants completed the reaction time and Stroop tests prior to the initiation of the interventions (Month 1), midway through (Months 4–5), and after the completion of the interventions (Month 10). The Wisconsin Card Sort Test was administered only twice (Month 1 and Month 10), to decrease the likelihood of learning the task and remembering the strategies. Each testing session lasted for approximately 90 min. The participants completed the reaction time tests first, followed by the Stroop, and finally the Wisconsin Card Sort Test.

The following four *reaction time* tests were administered: simple reaction time (SRT), 8-choice reaction time (8-CRT), 8-choice incompatible reaction time (8I-CRT), and Go/No-Go reaction time (GNG). For these tests, the participants were seated in front of a reaction time board so that they could easily reach each of the eight lights placed in a semi-circle. Using their dominant hand, they were to press the home button and, when one of the lights was illuminated (after a variable warning tone, from 800 ms to 2,000 ms), they were to respond as quickly as possible. For the SRT test, only the light right of center was visible and illuminated during the trial (all other lights were covered). For the 8-CRT and 8I-CRT tests, all lights were visible. For the 8-CRT test, the participants had to move to the illuminated light. For the 8I-CRT, they had to move to the light opposite to the one that was illuminated. For the GNG test, all lights except for the two center ones were covered. When the right light was illuminated, the participants had to move as quickly as possible (which was identical to the SRT test). However, when the left light was illuminated, they were to stay on the home button. Timing of the lights and measurement of reaction time were controlled by a computer interfaced to the reaction time board and appropriate software. Each test was explained and then the participants were given the opportunity to practice (five to eight trials) before the beginning of data collection. All participants completed 16 trials of each test except for the GNG, for which they completed 28 trials (12 inhibition trials in addition to 16 “go” trials). Reaction time was measured as the time from the “go” signal until the hand moved off the home button. Movement time was defined as the time from releasing the home button until pressing the illuminated button.

Three types of the computerized Stroop tests were administered (Word, Color, and Word–Color conflict). The participants were to respond as quickly as possible to the word or color block by pressing 1, 2, or 3 for blue, green, or red, respectively². The stimuli remained on the screen until the participant entered a response. For the Word–Color test, they were to respond to the color of the word rather than the word itself (e.g., when the word BLUE was presented in green letters, they were to respond by pressing 2 for green). The participants were instructed to respond as quickly as possible while trying to make as few errors as possible. To become familiar with the procedures and to learn to associate the color and key, participants practiced each condition for 45 s prior to data collection for that condition. One hundred trials were completed. Reaction time and number of errors were recorded (but, given the trade-off between these variables, only reaction time was treated as a dependent variable).

The Wisconsin Card Sort Test requires executive processing but is not a speeded task. The participants were presented with a card and they were instructed to match it to one of three other cards that were presented. However, the rules were not explained. The participants were only informed whether their choices were correct or incorrect. Therefore, they had to figure out which rule was being used for the correct match. Moreover, at undisclosed points during the test, the rule changed and the participants again had to figure out which rule was now being used. The test was administered and a computer program evaluated performance. The number of categories completed, the number of total errors, and the number of perseverative errors was recorded.

Data Reduction and Analysis

For SRT, 8-CRT, and 8I-CRT, of the 16 trials that were given, the fastest and slowest reaction times were dropped and the remaining 14 trials for each test were averaged, to improve reliability. For the GNG, 28 trials were given (16 “go” and 12 “no-go”) and the reaction times for the “go” trials (with the highest and lowest reaction times removed before averaging) were analyzed. For the Stroop test, the reaction times for the Word, Color, and Word–Color conflict

²Note that the computerized version of the Stroop test, by requiring that participants learn to associate a particular key with a particular response, probably places higher demands on working memory than the traditional paper format, in which responses are given verbally.

tests were analyzed. The reaction times for the erroneous responses were removed before averaging. For the Wisconsin Card Sort Test number of errors and categories completed were totaled.

Before statistical analysis, dependent variables were log-transformed to address deviations from distributional normality, which is a common problem in reaction time measures. The effect of the intervention on performance in the neurocognitive tests was examined by conducting two (CARDIO vs. FLEX-TONE) by three (Month 1, Months 4–5, and Month 10) or, in the case of the Wisconsin Card Sort Test, two by two (Month 1 and Month 10) multivariate analyses of variance (MANOVAs) for each of the following four sets of dependent variables: (a) reaction times for SRT, 8-CRT, 8I-CRT, and GNG; (b) reaction times for Stroop Word, Stroop Color, and Stroop Word–Color conflict; (c) number of categories completed, total errors, and perseverative errors for the Wisconsin Card Sort Test; and (d) errors from GNG, Stroop Word, Stroop Color, and Stroop Word–Color conflict. Significant group by time interactions were followed-up by repeated-measures univariate analyses of variance per group and Bonferroni-corrected pairwise comparisons. Effect sizes are also reported [$d=(M_i-M_j)/SD_{\text{pooled}}$], to provide an estimate of the magnitude of differences.

The “cardiovascular fitness” hypothesis can be conceptualized as a causal chain. In this mediational model, the causal effect of an exercise intervention (the independent variable) on changes in neurocognitive function (the dependent variable) can be “explained” by a third variable (the mediator), namely the exercise-induced changes in aerobic fitness ($VO_2\text{peak}$). Mediation is analyzed by establishing whether the following four criteria are met [48–50]: (a) the intervention has a significant effect on the outcome variable(s); (b) the intervention has a significant effect on the theorized mediator; (c) the mediator is significantly related to the outcome variable(s); and (d) the significant effect of the intervention on the outcome variable (s) disappears once one controls for the theorized mediator. In the present study, we considered whether (a) the intervention had any significant effects on neurocognitive variables based on the aforementioned analyses of variance, (b) the intervention had a significant effect on $VO_2\text{peak}$; (c) the changes in $VO_2\text{peak}$ during the intervention period were related to changes in neurocognitive function during the same period; and (d) any significant intervention effects on neurocognitive function remained after controlling for changes in $VO_2\text{peak}$.

Results

Preliminary Analyses

There were no significant differences between the groups in terms of age or years of education. A two (group: CARDIO, FLEX-TONE) by two (time: Month 1, Month 10) ANOVA on $VO_2\text{peak}$ showed only a significant main effect of time, $F(1, 55)=38.060, p=0.001$. Although the participants in the CARDIO group had a somewhat larger increase in $VO_2\text{peak}$ (18%) than those in the FLEX-TONE group (13%), the interaction was not statistically significant (see Table 1). There were no between-group differences in $VO_2\text{peak}$ either at the beginning or at the end of the interventions.

Reaction Time Tests

A two (group: CARDIO, FLEX-TONE) by three (time: Month 1, Months 4–5, Month 10) MANOVA on reaction times from the SRT, 8-CRT, 8I-CRT, and GNG showed only a significant main effect of time, $F(4, 108)=2.686, p=0.035$. However, follow-up univariate analyses failed to find significant changes over time for any of these four variables (see Table 2).

Stroop Tests

A similar two by three (group by time) MANOVA was conducted on the reaction times for Stroop Word, Stroop Color, and Stroop Word–Color conflict. This analysis revealed a significant interaction, $F(3,109)=3.079$, $p=0.031$. Follow-up univariate analyses indicated a significant group by time interaction only for Word–Color conflict, $F(2,110)=4.185$, $p=0.018$. A follow-up ANOVA for the FLEX-TONE group showed a significant effect of time, $F(2,56)=4.365$, $p=0.017$. However, Bonferroni-corrected pairwise comparisons indicated only a trend for a decrease from Month 1 to Months 4–5 ($p=.056$, $d=.47$). A similar follow-up ANOVA for the CARDIO group also showed a significant effect of time, $F(2,54)=10.730$, $p=0.001$. The pairwise comparisons indicated a significant and large decrease from Month 1 to Month 10 ($p=0.001$, $d=0.93$). Thus, the analyses of change for the Stroop data showed a significant over-time decrease in reaction times only for the Word–Color conflict test and only within the CARDIO group (see Table 2).

Wisconsin Card Sort Test

A two (group: CARDIO, FLEX-TONE) by two (time: Month 1, Month 10) MANOVA on the number of categories completed, total errors, and perseverative errors from the Wisconsin Card Sort Test revealed no significant main effect of time ($p=0.104$) or a significant interaction ($p=0.393$).

Errors

A two by three (group by time) MANOVA on errors from GNG, Stroop Word, Stroop Color, and Stroop Word–Color conflict showed a significant interaction, $F(2,110)=3.571$, $p=0.031$. The follow-up univariate analyses revealed a significant interaction only for the errors associated with the Stroop Word–Color conflict task, $F(2,110)=7.82$, $p=0.048$. Within the FLEX-TONE group, a significant effect of time, $F(2,56)=3.247$, $p=0.046$, was accounted for only by a significant *increase* in errors from Months 4–5 to Month 10 ($p=.028$, $d=.52$). On the other hand, within the CARDIO group, a significant effect of time, $F(2,54)=7.958$, $p=0.001$, was attributable to significant decreases from Month 1 to Months 4–5 ($p=.007$, $d=.64$) and from Month 1 to Month 10 ($p=0.002$, $d=0.72$). There were no significant between-group differences at any time point.

Examination of the “Cardiovascular Fitness” Hypothesis

As noted in the previous sections, the intervention was effective, as evidenced by significant and large decreases in both reaction time and errors in the Stroop Word–Color conflict task among the participants in the CARDIO group. However, the two arms of the intervention (CARDIO, FLEX-TONE) did not have a differential effect on VO_2 peak (the group by time interaction was not significant, indicating similar improvements in aerobic fitness in both groups). This finding alone suffices to discount the possibility that the differential effects on neurocognitive function (i.e., finding significant effects only for the CARDIO group) could be explained by the mediational role of aerobic fitness.

In the interest of providing a more comprehensive analysis, we also examined whether there was any relationship between changes in VO_2 peak over the period of the intervention with changes in neurocognitive function over the same period. However, hierarchical regression analyses, controlling for age, gender, and years of education, revealed no significant relationships.

Discussion

The primary purpose of this study was to provide the first independent test of the “selective improvement” hypothesis and, as such, the first attempt to replicate the Kramer et al. [12] findings. Consistent with this hypothesis, exercise training (CARDIO or FLEX-TONE) did not influence performance in tasks requiring relatively little executive control (i.e., SRT, Stroop Word, and Stroop Color). Statistically reliable improvements were limited to a task that requires a substantial involvement of executive control, namely the Stroop Word–Color conflict test. Performance on this test requires remembering a rule, inhibiting the initial impulse to respond, applying the rule, and ultimately making a decision that is incompatible with the initial impulse. Only the CARDIO group showed a significant improvement over time, both by lowering the reaction time and by decreasing the number of errors. These improvements were not only statistically reliable but also of large magnitude (i.e., effect sizes of 0.93 for reaction time and 0.72 for errors from Month 1 to Month 10) and, therefore, of possible practical meaningfulness. On the contrary, the FLEX-TONE group showed a considerably smaller and non-significant decrease in reaction time and an inconsistent pattern of change in errors (an initial non-significant decrease followed by a significant increase).

Nevertheless, support for the “selective improvement hypothesis” could be characterized as partial due to the fact that other tasks requiring some (i.e., 8-CRT, 8I-CRT) or, presumably, substantial (i.e., GNG) involvement of executive control did not show significant group by time interaction effects. In particular, the difference between the results obtained from the Stroop Word–Color conflict test and the GNG, both of which were theorized to belong to the “executive control” category of Colcombe and Kramer's [35] four-category conceptual model (i.e., speed, visuospatial, controlled, executive) might seem puzzling.

Since the exact cognitive requirements of these two tasks differ, there could be several explanations for the differential effects of exercise. We consider two as perhaps more likely. First, functional brain imaging studies have shown that both overlapping and distinct frontal areas are involved in the performance of these tasks. On the one hand, the anterior cingulate cortex, which is presumed to play a role in attentional control and error detection, has been found to be activated during both the GNG and the Stroop [51]. On the other hand, performance of the GNG task appears to involve an area extending from the right dorsolateral prefrontal cortex to the inferior frontal gyrus and the insula [51–53]. In contrast, the Stroop task appears to be strongly left-lateralized, perhaps due to its greater reliance on language. Performance on this task is associated mainly with activation of the left dorsolateral prefrontal cortex and the middle frontal gyrus [51,54]³. Thus, it is possible that the effects of aerobic exercise are lateralized [55] or localized to areas specific to the Stroop Word–Color conflict task. Interestingly, after a 6-month aerobic exercise intervention, participants were found to have increases in brain volume in the anterior cingulate cortex and middle frontal gyrus [56], areas found to be involved during the Stroop task.

Second, because of the highly automated nature of reading in adults, the incorrect response in the Stroop Word–Color conflict task is strongly prepotentiated. As a result, the demands that the Stroop task places on cognitive processes involved in attention and interference resolution are comparatively greater than those involved in the GNG. Combined with the requirement to respond under time pressure, this highly demanding nature of the Stroop Word–Color task is precisely the reason why it is also commonly used as a stress-induction procedure in laboratory studies [57]. As other authors have noted, aerobic exercise is more likely to benefit performance on tasks that involve higher, rather than lower, cognitive demands [24,58,59].

³We consider the results of imaging studies of the Stroop as applicable to the present study because, like the present study, imaging studies typically also involve the computerized version of the test and responses are entered through a keypad.

The Wisconsin Card Sort Test shares with the Stroop task the common features of set-switching and response suppression. Accordingly, imaging studies reveal some commonality in the frontal areas involved. For example, both the Wisconsin Card Sort Test and the Stroop activate the dorsolateral prefrontal cortex and the anterior cingulate cortex, although the former appears to do so bilaterally (but see Ref. [60]), whereas the latter is left-lateralized [52,54,61]. At the same time, the Wisconsin Card Sort Test involves some unique areas, both frontal (e.g., ventromedial and orbitofrontal cortices) and non-frontal (e.g., the inferior parietal lobule bilaterally, the basal ganglia, and the cerebellum), and an apparently wider distributed network. Although these dissimilarities in neural substrates could again account for the differential effects of aerobic exercise, a simpler explanation might be that the Wisconsin Card Sort Test was not speeded. Consistent with the aforementioned postulate that exercise is more likely to have beneficial effects on tasks with higher cognitive demands, it has been argued that exercise is also more likely to have beneficial effects on speeded, rather than non-speeded, tasks [58, 59]. This could be conceived as adding an important “boundary condition” to the “selective improvement” hypothesis, suggesting that it should not be presumed to apply to non-timed or non-speeded executive-control tasks.

The second purpose of the present study was to test the “cardiovascular fitness” hypothesis. Although the CARDIO intervention was effective in improving performance in the Stroop Word–Color conflict task, there was no evidence that this beneficial effect was mediated by changes in aerobic fitness. This conclusion is substantiated by two pieces of evidence. First, significant decreases in reaction time and errors in the Stroop Word–Color conflict task were found only in the CARDIO group but significant and substantial improvements in VO_{2peak} (by 18% and 13% in CARDIO and FLEX-TONE, respectively) were found in both groups, with no significant difference between them. If fitness gains mediated the effects of exercise on neurocognitive function, then the FLEX-TONE group should have also experienced benefits in neurocognitive function. Second, there was no evidence of a relationship between fitness gains and changes in neurocognitive function. Collectively, the data from the present study provided no support for the “cardiovascular fitness” hypothesis.

This finding adds one more piece of evidence to the growing data base supporting the ongoing paradigmatic shift in this area of research. As noted in the introduction, neither the meta-analysis of 18 intervention trials by Colcombe and Kramer [35] nor the meta-regression analysis of 37 studies (571 effect sizes) by Etnier et al. [39,50] have yielded any support for the idea that gains in aerobic fitness are related to changes in neurocognitive function. Collectively, these findings are instigating an evolution of the conceptual paradigm, shifting research attention to mediators other than aerobic fitness. After years of focusing on the role of fitness [36–38], Kramer and Erickson [62] have recently acknowledged that there seems to be “little evidence of a significant relationship between fitness change and cognitive change” (p. 343). Similarly, according to Etnier et al. [39], “our focus on aerobic fitness may have been misguided” (p. 126). Instrumental in this paradigmatic shift is the emergence of new information on other biological mediators by which exercise could benefit neurocognitive function, including brain vascularization, neurogenesis, and upregulation of neurotrophic factors (e.g., brain-derived neurotrophic factor or BDNF). None of these seem to be dependent on aerobic fitness. For example, high-volume or strenuous running, which can stimulate great gains in aerobic fitness, fails to increase neurogenesis beyond a plateau, does not benefit learning [63], and can have detrimental effects on BDNF [64,65]. In a human brain imaging study, it was reported that “fitness itself has little effect on brain tissue density” (p. 178) although it was found to moderate age-related declines in tissue density [66].

It should be noted that absolute levels of aerobic fitness (not changes in fitness induced by an exercise intervention) have been found to correlate positively with neurocognitive function. For example, in the Etnier et al. meta-regression analysis [39,50], 32 of 37 correlations reported

in 10 cross-sectional studies indicated a positive relationship between aerobic fitness and neurocognitive task performance. However, on average, fitness explained only 8% of the variance in cognition. In the present study, higher absolute levels of VO_2peak were also associated with better neurocognitive task performance across a range of tasks. However, similar to the results of the meta-regression analysis, VO_2peak typically accounted for less than 10% of the variance in neurocognitive task performance. These findings can probably be attributed to common “third” factors that influence overall health and vitality (of which fitness and cognition could both be markers), such as genetics or lifestyle characteristics.

The main strengths of the present study are the long duration of the intervention (leading to robust fitness gains), the assessment of neurocognitive function at three points in time, and the hypothesis-driven selection of dependent variables. On the other hand, the study also had several noteworthy limitations, which readers should take into account in interpreting its findings. First, since this study was built within a larger RCT, a number of participants who otherwise might have been included were excluded due to the screening criteria of the parent study (i.e., individuals with a history of cancer or those taking medication that could affect their immune function). This element could have had a negative impact on the representativeness of the sample and, thus, the generalizability of the findings to the population of older adults. Also relevant to this issue is the fact that 26% of the participants in the larger RCT declined the request to take part in neurocognitive testing since it entailed an additional commitment of time.

Second, the sample included 41 women but only 16 men, resulting in an unequal (72% to 28%) gender representation. It is possible that the preponderance of women could have influenced the results. For example, in the meta-analysis of Colcombe and Kramer [35], studies with samples consisting of more women than men yielded a much higher average effect size for the effects of exercise interventions on cognitive function (0.604) than studies with samples consisting of more men than women (0.150). It seems that this gender difference is due to the apparently neuroprotective effects of estrogen or estrogen replacement therapy [67] and an exercise–estrogen interaction [68–70]. However, the elderly women in the present study were all postmenopausal and only five of the 21 in CARDIO and two of the 20 in FLEX-TONE were receiving hormone replacement therapy. Therefore, the unequal gender representation notwithstanding, the influence of the preponderance of women on the results was probably attenuated.

Third, it is possible that the failure to find statistically significant interaction effects for additional tasks that require some degree of executive control (e.g., 8I-CRT, GNG, or the Wisconsin Card Sort Test), besides the Stroop Word–Color conflict test, might have been due to the limited statistical power afforded by the sample size of 57. However, examination of effect sizes (variance in the dependent variable accounted for by the group by time interaction) indicated that interaction effects were generally weak and, therefore, probably of limited practical significance. Moreover, a much larger sample size would have been required to attain statistical significance.

Additional study limitations included the estimation of aerobic fitness by a treadmill protocol rather than its direct assessment by analysis of expired gases, the recruitment of apparently high-functioning volunteers, and the absence of genetic (e.g., see [71]) and brain imaging data (e.g., see [66]). Future studies should be designed to address these limitations.

In summary, this study provided partial support for the “selective improvement” hypothesis and found no support for the “cardiovascular fitness” hypothesis. Based on the findings reported here, it could be concluded that the effects of aerobic exercise training for up to 10 months on indices of neurocognitive function do not seem to be generalized but appear, instead,

to be limited to *speeded* tasks toward the *high end* of the executive-control continuum. Gains in aerobic fitness do not appear to be related to changes in neurocognitive function.

Acknowledgments

This work was supported by a grant from the National Institute of Allergy and Infectious Diseases (R01 AI49956) to M.L.K. and a University Research Grant to A.L.S.-O.

References

1. Hedden T, Gabrieli JDE. Insights into the ageing mind: A view from cognitive neuroscience. *Nat Rev Neurosci* 2004;5:87–96. [PubMed: 14735112]
2. Park DC, Polk TA, Mikels JA, Taylor SF, Marshuetz C. Cerebral aging: Integration of brain and behavioral models of cognitive function. *Dialogues in Clinical Neuroscience* 2001;3:151–165.
3. Barnes DE, Yaffe K, Satariano WA, Tager IB. A longitudinal study of cardiorespiratory fitness and cognitive function in healthy older adults. *J Am Geriatr Soc* 2003;51:459–465. [PubMed: 12657064]
4. Laurin D, Verreault R, Lindsay J, MacPherson K, Rockwood K. Physical activity and risk of cognitive impairment and dementia in elderly persons. *Arch Neurol* 2001;58:498–504. [PubMed: 11255456]
5. Lytle ME, Vander Bilt J, Pandav RS, Dodge HH, Ganguli M. Exercise level and cognitive decline: The MoVIES project. *Alzheimer Dis Assoc Disord* 2004;18:57–64. [PubMed: 15249848]
6. Van Gelder BM, Tjihuis MAR, Kalmijn S, et al. Physical activity in relation to cognitive decline in elderly men: The FINE study. *Neurology* 2004;63:2316–2321. [PubMed: 15623693]
7. Weuve J, Kang JH, Manson JAE, Breteler MMB, Ware JH, Grodstein F. Physical activity, including walking, and cognitive function in older women. *J Am Med Assoc* 2004;292:1454–1461.
8. Yaffe K, Barnes D, Nevitt M, Lui LY, Covinsky K. A prospective study of physical activity and cognitive decline in elderly women: Women who walk. *Arch Intern Med* 2001;161:1703–1708. [PubMed: 11485502]
9. Hillman CH, Motl RW, Pontifex MB, et al. Physical activity and cognitive function in a cross-section of younger and older community-dwelling individuals. *Health Psychol* 2006;25:678–687. [PubMed: 17100496]
10. Van Boxtel MPJ, Paas FGWC, Houx PJ, et al. Aerobic capacity and cognitive performance in a cross-sectional aging study. *Med Sci Sports Exerc* 1997;29:1357–1365. [PubMed: 9346168]
11. Fabre C, Chamari K, Mucci P, Massé-Biron J, Préfaut C. Improvement of cognitive function by mental and/or individualized aerobic training in healthy elderly subjects. *Int J Sports Med* 2002;23:415–421. [PubMed: 12215960]
12. Kramer AF, Hahn S, Cohen NJ, et al. Ageing, fitness and neurocognitive function. *Nature* 1999;400:418–419. [PubMed: 10440369]
13. Black JE, Isaacs KR, Anderson BJ, Alcantara AA, Greenough WT. Learning causes synaptogenesis, whereas motor activity causes angiogenesis, in cerebellar cortex of adult rats. *Proc Natl Acad Sci* 1990;87:5568–5572. [PubMed: 1695380]
14. Isaacs KR, Anderson BJ, Alcantara AA, Black JE, Greenough WT. Exercise and the brain: Angiogenesis in the adult rat cerebellum after vigorous physical activity and motor skill learning. *J Cereb Blood Flow Metab* 1992;12:110–119. [PubMed: 1370068]
15. Swain RA, Harris AB, Wiener EC, et al. Prolonged exercise induces angiogenesis and increases cerebral blood volume in primary motor cortex of the rat. *Neuroscience* 2003;117:1037–1046. [PubMed: 12654355]
16. van Praag H, Kempermann G, Gage FH. Running increases cell proliferation and neurogenesis in the adult mouse dentate gyrus. *Nat Neurosci* 1999;2:266–270. [PubMed: 10195220]
17. van Praag H, Shubert T, Zhao C, Gage FH. Exercise enhances learning and hippocampal neurogenesis in aged mice. *J Neurosci* 2005;25:8680–8685. [PubMed: 16177036]
18. Berchtold NC, Kessler JP, Cotman CW. Hippocampal brain-derived neurotrophic factor gene regulation by exercise and the medial septum. *J Neurosci Res* 2002;68:511–521. [PubMed: 12111841]

19. Gómez-Pinilla F, So V, Kesslak JP. Spatial learning and physical activity contribute to the induction of fibroblast growth factor: Neural substrates for increased cognition associated with exercise. *Neuroscience* 1998;85:53–61. [PubMed: 9607702]
20. Trejo JL, Carro E, Torres-Alemán I. Circulating insulin-like growth factor I mediates exercise-induced increases in the number of new neurons in the adult hippocampus. *J Neurosci* 2001;21:1628–1634. [PubMed: 11222653]
21. Vaynman S, Ying Z, Gómez-Pinilla F. Hippocampal BDNF mediates the efficacy of exercise on synaptic plasticity and cognition. *Eur J Neurosci* 2004;20:2580–2590. [PubMed: 15548201]
22. Pereira AC, Huddleston DE, Brickman AM, et al. An in vivo correlate of exercise-induced neurogenesis in the adult dentate gyrus. *Proc Natl Acad Sci* 2007;104:5638–5643. [PubMed: 17374720]
23. Dustman RE, Ruhling RO, Russell EM, et al. Aerobic exercise training and improved neuropsychological function of older individuals. *Neurobiol Aging* 1984;5:35–42. [PubMed: 6738784]
24. Hawkins HL, Kramer AF, Capaldi D. Aging, exercise, and attention. *Psychol Aging* 1992;7:643–653. [PubMed: 1466833]
25. Rikli RE, Edwards DJ. Effects of a three-year exercise program on motor function and cognitive processing speed in older women. *Res Q Exerc Sport* 1991;62:61–67. [PubMed: 2028094]
26. Williams P, Lord SR. Effects of group exercise on cognitive functioning and mood in older women. *Aust New Zeal J Publ Health* 1997;21:45–52.
27. Blumenthal JA, Emery CF, Madden DJ, et al. Long-term effects of exercise on psychological functioning in older men and women. *J Gerontol* 1991;46:P352–P361. [PubMed: 1940092]
28. Hassmén P, Ceci R, Bäckman L. Exercise for older women: A training method and its influences on physical and cognitive performance. *Eur J Appl Physiol* 1992;64:460–466.
29. Hill RD, Storandt M, Malley M. The impact of long-term exercise training on psychological function in older adults. *J Gerontol* 1993;48:P12–P17. [PubMed: 8418145]
30. Madden DJ, Blumenthal JA, Allen PA, Emery CF. Improving aerobic capacity in healthy older adults does not necessarily lead to improved cognitive performance. *Psychol Aging* 1989;4:307–320. [PubMed: 2803624]
31. Panton LB, Graves JE, Pollock ML, Hagberg JM, Chen W. Effect of aerobic and resistance training on fractionated reaction time and speed of movement. *J Gerontol* 1990;45:M26–M31. [PubMed: 2295775]
32. Dustman RE, Emmerson R, Shearer D. Physical activity, age, and cognitive neuropsychological function. *J Aging Phys Act* 1994;2:143–181.
33. West RL. An application of the prefrontal cortex function theory to cognitive aging. *Psychol Bull* 1996;120:272–292. [PubMed: 8831298]
34. West R. In defense of the frontal lobe hypothesis of cognitive aging. *J Int Neuropsychol Soc* 2000;6:727–729. [PubMed: 11011518]
35. Colcombe S, Kramer AF. Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychol Sci* 2003;14:125–130. [PubMed: 12661673]
36. Kramer AF, Erickson KI, Colcombe SJ. Exercise, cognition, and the aging brain. *J Appl Physiol* 2006;101:1237–1242. [PubMed: 16778001]
37. Kramer AF, Hahn S, McAuley E. Influence of aerobic fitness on the neurocognitive function of older adults. *J Aging Phys Act* 2000;8:379–385.
38. Kramer AF, Colcombe S, Erickson K, et al. Effects of aerobic fitness training on human cortical function: A proposal. *J Mol Neurosci* 2002;19:227–231. [PubMed: 12212786]
39. Etnier JL, Nowell PM, Landers DM, Sibley BA. A meta-regression to examine the relationship between aerobic fitness and cognitive performance. *Brain Res Rev* 2006;52:119–130. [PubMed: 16490256]
40. Kohut M, McCann DA, Russell DW, et al. Aerobic exercise, but not flexibility/resistance exercise, reduces serum IL-18, CRP, and IL-6, independent of β -blockers, BMI, and psychosocial factors in older adults. *Brain Behav Immun* 2006;20:201–209. [PubMed: 16504463]

41. American College of Sports Medicine. ACSM's Guidelines for Exercise Testing and Prescription. Vol. 7th ed. Lippincott, Williams & Wilkins; Philadelphia, PA: 2006.
42. Rikli RE, Jones CJ. Functional fitness normative scores for community-residing older adults, ages 60–94. *J Aging Phys Act* 1999;7:162–181.
43. Rikli, RE.; Jones, CJ. *Senior Fitness Test Manual*. Human Kinetics; Champaign, IL: 2001.
44. Sheridan LK, Fitzgerald HE, Adams KM, et al. Normative Symbol Digit Modalities Test performance in a community-based sample. *Arch Clin Neuropsychol* 2006;21:23–28. [PubMed: 16139470]
45. Smith, A. *Symbol Digit Modalities Test manual, Revised 1982*. Western Psychological Services; Los Angeles, CA: 1982.
46. Foster C, Jackson AS, Pollock ML, et al. Generalized equations for predicting functional capacity from treadmill performance. *Am Heart J* 1984;107:1229–1234. [PubMed: 6720550]
47. Rikli RE, Jones CJ. The reliability and validity of a 6-minute walk test as a measure of physical endurance in older adults. *J Aging Phys Act* 1998;6:363–375.
48. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986;51:1173–1182. [PubMed: 3806354]
49. Judd CM, Kenny DA. Process analysis: Estimating mediation in treatment evaluations. *Eval Rev* 1981;5:602–619.
50. Etnier, J. Interrelationships of exercise, mediator variables, and cognition.. In: Spirduso, WW.; Poon, LW.; Chodzko-Zajko, W., editors. *Exercise and its Mediating Effects on Cognition*. Human Kinetics; Champaign, IL: 2008. p. 13-30.
51. Nee DE, Wager TD, Jonides J. Interference resolution: Insights from a meta-analysis of neuroimaging tasks. *Cognit Affect Behav Neurosci* 2007;7:1–17. [PubMed: 17598730]
52. Buchsbaum BR, Greer S, Chang WL, Berman KF. Meta-analysis of neuroimaging studies of the Wisconsin Card-Sorting Task and component processes. *Hum Brain Mapp* 2005;25:35–45. [PubMed: 15846821]
53. Simmonds DJ, Pekar JJ, Mostofsky SH. Meta-analysis of Go/No-Go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia* 2008;46:478–493.
54. Alvarez JA, Emory E. Executive function and the frontal lobes: A meta-analytic review. *Neuropsychol Rev* 2006;16:17–42. [PubMed: 16794878]
55. Hall EE, Petruzzello SJ. Frontal asymmetry, dispositional affect, and physical activity in older adults. *J Aging Phys Act* 1999;7:76–90.
56. Colcombe SJ, Erickson KI, Scalf PE, et al. Aerobic exercise training increases brain volume in aging humans. *J Gerontol Med Sci* 2006;61A:1166–1170.
57. Renaud P, Blondin JP. The stress of Stroop performance: Physiological and emotional responses to color-word interference, task pacing, and pacing speed. *Int J Psychophysiol* 1997;27:87–97. [PubMed: 9342640]
58. Chodzko-Zajko, W. National blueprint: Increasing physical activity among adults 50 and older: Implications for future physical activity and cognitive functioning research.. In: Poon, LW.; Chodzko-Zajko, W.; Tomporowski, RD., editors. *Active Living, Cognitive Functioning, and Aging*. Human Kinetics; Champaign, IL: 2006. p. 1-14.
59. Chodzko-Zajko W, Moore KA. Physical fitness and cognitive functioning in aging. *Exerc Sports Sci Rev* 1994;22:195–220.
60. Lie CH, Specht K, Marshall JC, Fink GR. Using fMRI to decompose the neural processes underlying the Wisconsin Card Sorting Test. *NeuroImage* 2006;30:1038–1049. [PubMed: 16414280]
61. Derrfuss J, Brass M, Neumann J, von Cramon DY. Involvement of the inferior frontal junction in cognitive control: Meta-analyses of switching and Stroop studies. *Hum Brain Mapp* 2005;25:22–34. [PubMed: 15846824]
62. Kramer AF, Erickson KI. Capitalizing on cortical plasticity: Influence of physical activity on cognition and brain function. *Trends Cogn Sci* 2007;11:342–348. [PubMed: 17629545]

63. Rhodes JS, van Praag H, Jeffrey S, et al. Exercise increases hippocampal neurogenesis to high levels but does not improve spatial learning in mice bred for increased voluntary wheel running. *Behav Neurosci* 2003;117:1006–1016. [PubMed: 14570550]
64. Aguiar AS Jr, Tuon T, Pinho CA, et al. Intense exercise induces mitochondrial dysfunction in mice brain. *Neurochem Res* 2008;33:51–58. [PubMed: 17619145]
65. Soya H, Nakamura T, Deocaris CC, et al. BDNF induction with mild exercise in the rat hippocampus. *Biochem Biophys Res Commun* 2007;358:961–967. [PubMed: 17524360]
66. Colcombe SJ, Erickson KI, Raz N, et al. Aerobic fitness reduces brain tissue loss in aging humans. *J Gerontol Med Sci* 2003;58A:176–180.
67. Garcia-Segura LM, Azcoitia I, DonCarlos LL. Neuroprotection by estradiol. *Prog Neurobiol* 2001;63:29–60. [PubMed: 11040417]
68. Berchtold NC, Kesslack JP, Pike CJ, Adlard PA, Cotman CW. Estrogen and exercise interact to regulate brain-derived neurotrophic factor mRNA and protein expression in the hippocampus. *Eur J Neurosci* 2001;14:1992–2002. [PubMed: 11860494]
69. Cotman CW, Berchtold NC. Exercise: A biobehavioral intervention to enhance brain health and plasticity. *Trends Neurosci* 2002;25:295–301. [PubMed: 12086747]
70. Erickson KI, Colcombe SJ, Elavsky S, et al. Interactive effects of fitness and hormone treatment on brain health in postmenopausal women. *Neurobiol Aging* 2007;28:179–185. [PubMed: 16406152]
71. Etnier JL, Caselli RJ, Reiman EM, et al. Cognitive performance in older women relative to ApoE- ϵ 4 genotype and aerobic fitness. *Med Sci Sports Exerc* 2007;39:199–207. [PubMed: 17218903]

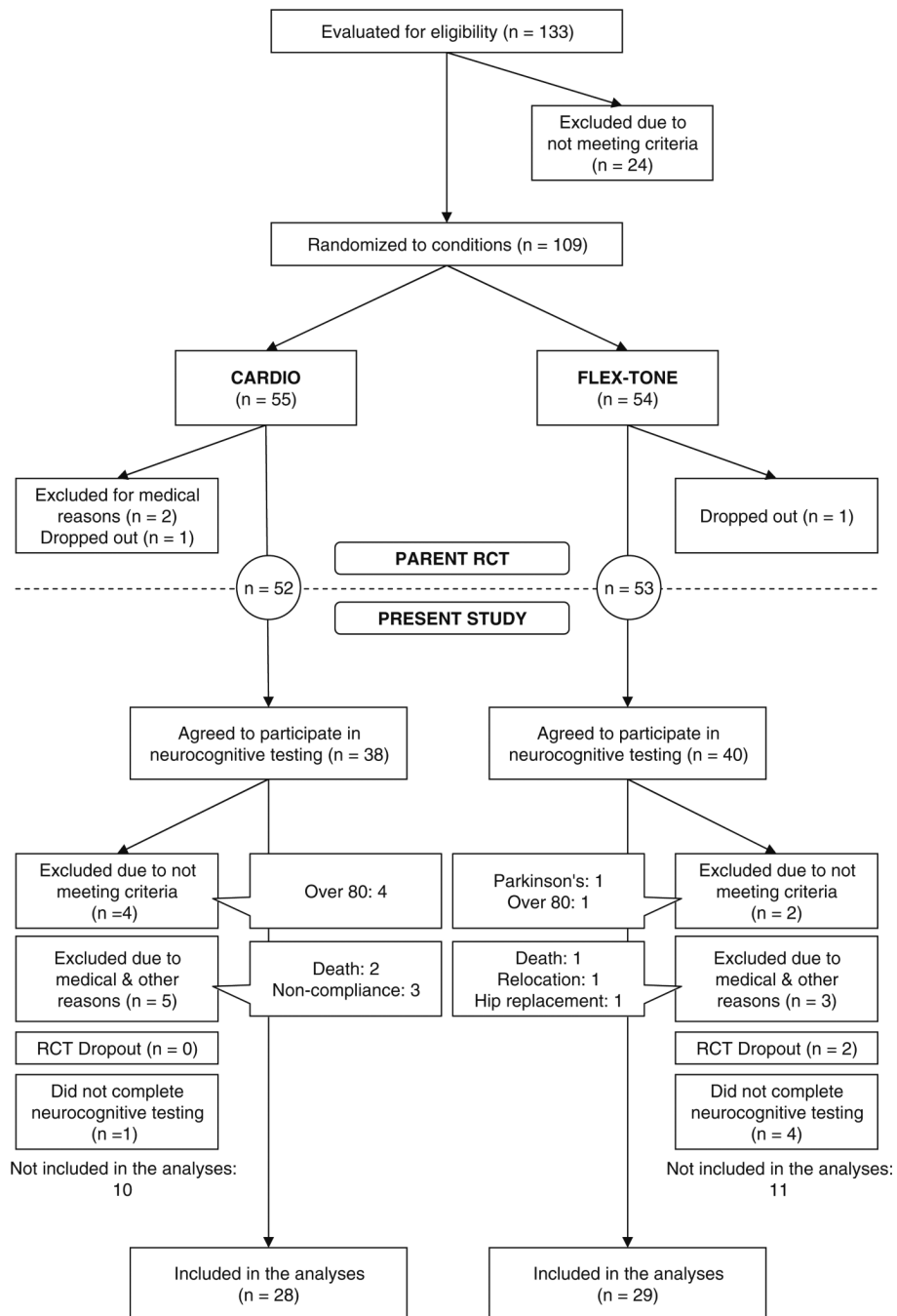


Fig. 1. Flow diagram showing the inclusion and exclusion of participants throughout the study

Table 1

Descriptive statistics (M±SD) for participant characteristics

	CARDIO Group			FLEX-TONE Group		
	Total (n=28) M±SD	Men (n=7) M±SD	Women (n=21) M±SD	Total (n=29) M±SD	Men (n=9) M±SD	Women (n=20) M±SD
Age (years)	69.86±4.59	68.55±2.65	70.30±5.06	70.52±4.47	69.80±5.13	70.85±4.24
Education (years)	15.82±2.83	17.57±3.36	15.24±2.45	14.79±2.34	16.00±2.4	14.25±2.15
VO ₂ peak at Month 1 (ml·kg ⁻¹ ·min ⁻¹)	29.03±7.81	34.40±7.48	27.23±7.19	28.41±9.94	34.30±9.33	25.76±9.23
VO ₂ peak at Month 10 (ml·kg ⁻¹ ·min ⁻¹)	34.32±6.23	37.71±4.39	33.18±6.43	32.12±9.66	39.82±9.02	28.65±7.88

Table 2

Descriptive statistics (M±SD) for all dependent variables

	CARDIO Group			FLEX-TONE Group		
	Month 1 M±SD	Months 4–5 M±SD	Month 10 M±SD	Month 1 M±SD	Months 4–5 M±SD	Month 10 M±SD
Simple RT (ms)	345±30	338±29	357±68	345±35	333±29	336±38
8-CRT (ms)	431±40	440±48	435±35	452±82	426±44	428±62
8I-CRT (ms)	806±220	806±228	767±229	801±196	780±170	764±204
GNG RT (ms)	460±64	477±88	460±64	460±74	457±70	442±72
GNG errors	1.00±1.39	0.71±1.08	0.44±0.73	1.07±1.07	0.72±0.84	0.77±0.82
Stroop Word RT (ms)	891±141	895±172	865±138	845±115	842±138	836±136
Stroop Word errors	3.96±5.83	2.71±3.40	2.29±3.62	1.62±1.37	2.16±1.80	1.78±1.46
Stroop Color RT (ms)	846±120	875±144	915±270	821±115	816±113	860±233
Stroop Color errors	2.64±3.19	2.57±2.59	9.86±10.33	1.55±1.55	1.84±1.39	8.39±10.09
Stroop Word–Color RT (ms)	1315±302	1275±296	1155±212	1263±294	1165±216	1192±268
Stroop Word–Color errors	6.43±5.59	3.75±4.53	3.38±3.91	3.72±3.25	2.52±2.78	3.32±2.64
WCST errors (total)	18.32±9.90		15.48±9.27	19.90±9.50		20.08±9.03
WCST errors (perseverative)	9.82±8.68		7.67±4.74	9.93±5.11		10.20±5.32
WCST categ. completed	3.11±1.55		3.14±1.76	2.72±1.49		2.56±1.55

RT Simple reaction time, CRT choice reaction time, I-CRT incompatible CRT, GNG Go/No-Go, WCST Wisconsin Card Sort Test