Research Article

FITNESS EFFECTS ON THE COGNITIVE FUNCTION OF OLDER ADULTS: A Meta-Analytic Study

Stanley Colcombe and Arthur F. Kramer

Beckman Institute and Department of Psychology, University of Illinois, Urbana

Abstract—A meta-analytic study was conducted to examine the hypothesis that aerobic fitness training enhances the cognitive vitality of healthy but sedentary older adults. Eighteen intervention studies published between 1966 and 2001 were entered into the analysis. Several theoretically and practically important results were obtained. Most important, fitness training was found to have robust but selective benefits for cognition, with the largest fitness-induced benefits occurring for executive-control processes. The magnitude of fitness effects on cognition was also moderated by a number of programmatic and methodological factors, including the length of the fitness-training intervention, the type of the intervention, the duration of training sessions, and the gender of the study participants. The results are discussed in terms of recent neuroscientific and psychological data that indicate cognitive and neural plasticity is maintained throughout the life span.

The study of the relationship between physical fitness, and more specifically aerobic fitness, and cognition dates back several decades to the pioneering research of Spirduso and her colleagues. For example, in a classic study, Spirduso and Clifford (1978) compared young and old racquet sportsmen, runners, and sedentary individuals on simple, choice, and movement time tasks. The older athletes' performance on these tasks was substantially better than the older sedentary adults', and indeed was similar to the performance of the young sedentary adults. Smaller fitness benefits were observed for young adults. Similar results, that is, better performance by older high-fit than older low-fit individuals, have been found in a large number of cross-sectional studies (see Dustman, Emmerson, & Shearer, 1994, and Etnier et al., 1997, for reviews of this literature).

The cross-sectional nature of these studies, however, complicates their interpretation. Thus, the positive effects of fitness on perceptual, cognitive, and motor processes may reflect a predisposition of the exercisers toward fast and accurate responding rather than a benefit of aerobic fitness achieved through exercise. A number of researchers have at least partially circumvented the problem of self-selection by employing longitudinal exercise interventions. However, the data obtained from longitudinal studies have been equivocal (Dustman et al., 1994). Clearly, there have been some notable successes in that aerobically trained individuals have outperformed nonaerobic control subjects on a variety of cognitive tasks (e.g., Dustman et al., 1984; Hawkins, Kramer, & Capaldi, 1992; Rikli & Edwards, 1991). However, other intervention studies have failed to observe such benefits to performance (e.g., Blumenthal et al., 1991; Hill, Storandt, & Malley, 1993). Thus, an important unanswered question concerns why some studies find improvements in performance with enhanced aerobic fitness while other studies have failed to observe such a relationship. This is the question that we addressed in the present study through the application of meta-analytic techniques to longitudinal studies of fitness effects on the cognition of older adults.

Research using animal models provides reasonable grounds to expect aerobic exercise to have a positive impact on human cognitive function through a variety of cellular and molecular mechanisms. For example, Black, Isaacs, Anderson, Alcantara, and Greenough (1990) found increases in capillary density in the cerebellum when rats exercised on a running wheel. Other researchers found that increased aerobic fitness, engendered by running, enhanced cortical high-affinity choline uptake and increased dopamine receptor density in the brains of old rats (Fordyce & Farrar, 1991), increased brain-derived neurotrophic factor (BDNF) gene expression in rats (Neeper, Gomez, Choi, & Cotman, 1995), and increased the number of new cells in the hippocampus of mice (van Praag, Kempermann, & Gage, 1999). It is reasonable to speculate that these cellular, molecular, and neurochemical changes observed in rats and mice in response to exercise interventions may underlie the improvements in perceptual, cognitive, and motor processes in adult humans.

Despite these impressive results with animals, the fitness interventions with humans, as we have indicated, have produced less reliable effects on performance in a variety of perceptual, cognitive, and motor tasks. Clearly, some of the ambiguity in the results obtained in the human studies could be the result of methodological factors. For example, studies have employed widely varying age groups. Studies have also differed with respect to the nature, intensity, and length of the aerobic fitness manipulations; the type of fitness measures employed; the general health and fitness level of the participants at the beginning of the study; subjects' gender; the tasks used to measure aspects of cognition; and the nature of the control groups.

However, studies of the influence of fitness training on cognition have differed not only with regard to these methodological concerns, but also with regard to the theoretical framework in which they have been conducted. Differences in theoretical framework, in turn, influence the tasks selected to index fitness effects. Many researchers have taken an atheoretical approach, selecting cognitive tasks with little or no theoretical justification. Other researchers have, either explicitly or implicitly, assumed that fitness effects would be most likely to be observed in tasks such as simple reaction time or finger tapping, which presumably tap low-level central nervous system function uncontaminated by subject strategies or high-level cognition (Dustman et al., 1984; Spirduso & Clifford, 1978). Hereafter, we refer to this proposal as the speed hypothesis. Other researchers have suggested that fitness effects might be most readily observed for visuospatial tasks (Shay & Roth, 1992; Stones & Kozma, 1989), because visuospatial processes have been demonstrated to be more susceptible to aging than verbal skills. Hereafter, we refer to this hypothesis as the visuospatial hypothesis. Chodzko-Zajko and his colleagues (Chodzko-Zajko, 1991;
Fitness and Cognition

Chodzko-Zajko & Moore (1994) have suggested tasks that require controlled, effortful processing should be more sensitive to fitness differences among older adults than tasks that can be executed via automatic processing. This controlled-processing hypothesis is based on earlier research by Schneider and Shiffrin (1977), who proposed a theory of skill acquisition in which task components or skills transition from controlled, effortful processing to automatic processing with consistent practice. Finally, we (Kramer, Hahn, et al., 1999; see also Hall, Smith, & Keele, 2001) have argued that improvements in fitness would be reflected in enhancements in executive-control processes such as coordination, inhibition, scheduling, planning, and working memory (Shallice, 1994). These tasks stand in contrast to controlled tasks in that they do not become automatic over time, and require constant mediation by a central executor. Executive-control processes and the brain areas that support them have been shown to be disproportionately sensitive to aging (West, 1996). Hereafter, we refer to this hypothesis as the executive-control hypothesis.

Although it is clear that these four hypotheses are at best crude approximations of theoretical accounts of fitness effects on human cognition, and that tasks and task components can only rarely be neatly and orthogonally decomposed into the processing components assumed in these hypotheses, we believe that these hypotheses provide at least a reasonable starting point in an attempt to make some sense of a literature that has often produced confusing and seemingly unreliable results. That is, even these incomplete and underspecified hypotheses offer the promise of providing some structure to and understanding of the database of studies that have investigated whether, and under what conditions, improvements in aerobic fitness result in enhanced cognition.

RESEARCH AIDS AND HYPOTHESES

In the present study, we quantitatively examined, in a meta-analysis, these four theoretical hypotheses as well as some methodological factors that may influence the extent to which enhancements in aerobic fitness result in improvements in cognition (Rosenthal, 1998). Meta-analysis is a particularly appropriate technique for this purpose because it enables one to summarize the relationship between two variables across different studies—in the present case, the effect of fitness interventions on cognition. However, in addition to yielding an overall effect size, as well as effect sizes for each study, the meta-analytic procedure enables one to determine whether one or more moderator variables (e.g., theoretical distinctions or methodological factors) influenced the outcome of the studies.

In several ways, we extended previous qualitative and quantitative analyses of the literature on the fitness-cognition relationship (Dustman et al., 1994; Etnier et al., 1997). First, we examined the theoretical proposals that have been made with regard to the processes expected to benefit most from enhancements in aerobic fitness. Second, we focused specifically on randomized fitness intervention trials that included control groups and on fitness training that extended from several months to several years. Third, we included studies conducted from 1966 to 2001 in our analyses. Finally, we focused on older adults, from 55 to 80 years of age. Identifying the interventions that have shown promise in enhancing the cognitive vitality of this large and growing segment of the population is particularly important given the demonstrated decline in a number of aspects of cognition across the adult life span.

METHOD

Literature Search

Our search efforts were primarily focused within the Current Contents, Education Research in Completion (ERIC), MedLine, PsychInfo, and PsychLit on-line databases. Periodic searches of these databases were conducted between October 2000 and July 2001 using combinations of the following keywords: age, ageing, age differences, older, elderly, exercise, fitness, cardiovascular, aerobic exercise, cognitive, cognition, and longitudinal. Additionally, we manually searched book chapters and peer-reviewed articles on the topic to locate potentially relevant studies. The on-line and manual searches yielded 167 articles, which were then reviewed to assess their appropriateness for this study. Articles were excluded if (a) the design was cross-sectional, (b) participants were not randomly assigned, (c) the exercise program was unsupervised, (d) the exercise program did not include an aerobic fitness component, or (e) participants’ average age was below 55. The remaining 18 articles yielded 197 effect sizes for further analyses. (These articles are included in our reference list here, marked with asterisks.)

Effect-Size Calculation

Primary analyses were conducted through the Comprehensive Meta-Analysis software package (BioStat, Englewood, New Jersey), using Hedges’s (1982) formula for calculating effect size: 

\[ g = (M_{post} - M_{pre})/SD_p, \]

where \( M_{post} \) is the preintervention mean task performance, \( M_{pre} \) is the postintervention mean, and \( SD_p \) is the pooled standard deviation. All effect sizes were weighted by Hedges’s sample-size correction: 

\[ c = 1 - 3/(4m - 9), \]

and \( m = 2N - 2 \). Effect sizes were coded such that positive numbers always reflected improvements in performance, and negative numbers reflected deterioration in performance.

Coding of Outcomes

Theoretical variables

Of primary interest in this study was the effect of exercise on cognitive processes identified by the four theoretical positions (speed, visuospatial, controlled processing, and executive control). Therefore, each task (or condition within a task) was categorized as to whether each of the four theoretical positions would predict that the task would be differentially sensitive to improvements in aerobic fitness. A task was coded as 1 for a given theoretical position if it should show differential improvement according to that position, and given a 0 otherwise.

Effect sizes were coded as belonging to the speed category if the task represented a measure of low-level neurological functioning, such as simple reaction time (e.g., the speed with which one can make a manual response to a flash of light) or finger-tapping speed. Likewise, effect sizes were coded under the visuospatial category if the task tapped the participants’ ability to transform or remember visual and spatial information (e.g., viewing three line drawings and then replicating them from memory as in the Benton Visual Retention Task). Effects coded to represent controlled processes were from tasks that require, at least initially, some cognitive control (e.g., pressing one key when presented with the letter C, but pressing a different key for the letter M, as in a choice reaction time task), but do not have those char-
acteristics exemplified by tasks in the executive category. Tasks in the executive category were those related to the planning, inhibition, and scheduling of mental procedures. For example, this category would include the Erickson flanker task, which requires participants to respond to a central cue (much as in a choice reaction time task), but simultaneously suppress a set of conflicting or irrelevant cues presented next to the target stimulus item.

Tasks in each of these categories ostensibly could be affected to some degree by components of tasks in other categories. And although traditionally many authors have argued that the four hypotheses are discrete theoretical positions, it might be more appropriate to view them as identifying partially discrete, but not mutually exclusive, pathways by which physical exercise might mediate cognitive functioning. As we have already suggested, a subcomponent of the flanker task (coded as an executive task) is much like the choice reaction time task (coded as a controlled process). Likewise, simple manual reaction time (coded as a speed measure) is surely a subcomponent of the choice reaction time task (coded as a controlled process). Given the difficulty of cleanly dividing most tasks according to the theoretical variables of interest, each cognitive task was coded with respect to each theoretical variable independently, and we allowed multiple category membership. Coding was independently performed by two researchers familiar with a broad range of task types; interrater reliability was .96.

**Participants’ characteristics**

The effects were also coded with respect to the average age of participants at the beginning of each intervention: young-old (55–65), middle-old (66–70), and old-old (71+). Additionally, although most studies did not report effects separately for males and females, we were able to assess the potential moderating effect of sex by evaluating the relative proportion of males and females in each study. Therefore, effects were also coded by the proportion of males and females: high male (≥50% male) and high female (>50% female).

Additionally, the target populations in the studies were most often community-dwelling, “normal” older adults, but a limited number of studies targeted clinical populations of one kind or another, ranging from depressed persons (Khatri et al., 2001) to geriatric mental patients (Palleschi, Vetta, De Gennaro, & Idone, 1994) and individuals with cardiopulmonary obstructive disorders (Emery, Schein, Hauck, & MacIntyre, 1998). Effect sizes drawn from studies in which the target population was in some way representative of a particular clinical population were coded as clinical, and others were coded as nonclinical. The relatively small number of studies examining clinical populations (4) did not allow a careful examination of each subgroup, so the clinical group was far from a homogeneous category. However, a preliminary evaluation of the effect of exercise on the cognitive functioning of these populations may prove useful.

**Characteristics of training interventions**

The training interventions involved a wide range of activities from walking to dancing and circuit training, but were quite easily divided into two groups: those that emphasized cardiovascular fitness in isolation (aerobic) and those that combined cardiovascular fitness training with strength training (combination). The studies were also coded with respect to the duration of the training session (short, 15–30 min; moderate, 31–45 min; and long, 46–60 min) and the length of the exercise intervention (short, 1–3 months; medium, 4–6 months; and long, 6+ months). Finally, each study was coded with respect to the relative amount of cardiovascular improvement shown in the participants, based on either estimated or actual VO²peak or max scores (unreported; moderate, 5–11%; and large, 12–25%).

**Analysis**

Each study reported multiple effect sizes, often spanning all four theoretical variables. All possible effect sizes were entered into the analysis of the effects of exercise on cognition generally, as well as analyses of the theoretical and moderator variables (Rosenthal & Rubin, 1986).

The first analysis conducted assessed the effects of exercise on cognitive performance generally, and included effect sizes derived from both experimental and control groups. Subsequent analysis of the effects of the theoretical and moderator variables was conducted separately for control and exercise groups (see Bangert-Drowns, 1986).

**RESULTS**

**Test of Heterogeneity**

Effect sizes were found to range from −0.9 to +6.4, and showed sufficient heterogeneity, Q(100) = 1.028, p < .00001, to justify further examination.

**Effects of Exercise**

Point estimates for effect size on all cognitive tasks were 0.164 (SE = 0.028, n = 96, p < .05) for control groups and 0.478 (SE = 0.029, n = 101, p < .01) for exercisers, suggesting that both groups improved between Times 1 and 2. However, the control groups’ improvement was about 1/8 a standard deviation, whereas the exercise groups’ improvement was nearly 1/2 a standard deviation, on average. Both values are significantly different from zero, and from each other (see Table 1).

**Exercisers**

**Theoretical variables**

An examination of the estimates for each task-process category revealed that exercise had the greatest effect on executive processes (g = 0.68, 1). Another option when coding the tasks would be to use a subtractive approach, in which, when possible, the more complex executive or controlled processes would be estimated by subtracting presumably lower-order processes (e.g., speed) from the higher-order ones. In this manner, we could presumably have reduced the contamination of basic speed of processing within controlled and executive processes. However, the logic of “pure insertion” required under such an approach is somewhat suspect, and, moreover, the extant data would not allow such a procedure to be uniformly applied to all studies or tasks. As a result, we opted for the more conservative, but uniform procedure described here.

2. Dividing the effects in this manner yielded 42 in the mostly male group, and 53 in the mostly female group. Although a finer-grained examination of these effects would be desirable (e.g., ±25%, 26–50%, 51–75%, 76–100%), the available data were not amenable to such analyses.
The meta-analysis revealed that when the participant population was more than half female, the group as a whole showed greater benefit than if the population was at least half male. Additionally, participants in the mid-old category seemed to benefit most from exercise. Interestingly, clinical and nonclinical populations showed similar improvements with exercise ($g = 0.47$ for clinical, $g = 0.48$ for nonclinical), suggesting that exercise interventions can be equally efficacious for clinical and nonclinical populations.

**Control Participants**

There were no reliable effects of theoretical variables on effect size for control participants, nor were there many reliable effects of the moderator variables. However, one finding is worth noting. In the control group, performance of the young-old ($g = 0.108, SE = 0.053, n = 23, p < .05$) and middle-old ($g = 0.258, SE = 0.045, n = 48, p < .05$) increased much more than performance of the old-old ($g = 0.076, SE = 0.058, n = 25, n.s.$).

**DISCUSSION**

We investigated two main hypotheses in this meta-analytic study. First, we examined whether aerobic fitness training can have a robust and beneficial influence on the cognition of sedentary older adults. The animal literature that has addressed this issue (e.g., Black et al., 1990; Nepper et al., 1995) suggests an affirmative answer, but a perusal of the literature on human aerobic training appears more equivocal. The answer provided by the present analysis is an unequivocal yes. Fitness training increased performance $0.5 SD$ on average, regardless of the type of cognitive task, the training method, or participants’ characteristics.

Although the present analysis establishes the efficacy of fitness training as a means to enhance the cognitive vitality of older adults, the examination of the moderator variables also begins to establish important boundary conditions on the relationship between fitness training and cognition, as well as to generate prescriptions for additional research.

Perhaps most important, the analysis suggests that robust but process-specific benefits accrue with fitness training. As we previously hypothesized (Kramer, Hahn, et al., 1999; see also Hall et al., 2001), executive-control processes showed the largest benefit of improved fitness. However, controlled processes (Schneider & Shiffrin, 1977), which at least partially overlap with executive processes, and visuospatial processes also showed reliable benefits from fitness training. All of these processes have shown large age-related performance decrements in previous studies and also appear to benefit from intellectual training (e.g., Kramer, Larish, Weber, & Bardell, 1999; Schaie & Willis, 1986). The present results, along with the extant animal literature, suggest that fitness training can also enhance cognitive vitality of older adults.

An interesting question for future research concerns the manner in which these performance changes are supported by changes in patterns of brain activation, as inferred from positron emission tomography, functional magnetic resonance imaging (fMRI), and optical imaging. Clearly, changes in cognitive performance must be mediated by changes in neural activation. However, it is unclear precisely what role cardiovascular fitness might play in instantiating these changes. A number of studies of aging, cognition, and brain function have found evidence for dedifferentiation, that is, less specificity among older

---

### Table 1. Results for significant moderating variables

<table>
<thead>
<tr>
<th>Moderator variable</th>
<th>Effect size</th>
<th>SE</th>
<th>n</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>0.164</td>
<td>0.028</td>
<td>96</td>
<td>*</td>
</tr>
<tr>
<td>Control</td>
<td>0.478</td>
<td>0.029</td>
<td>101</td>
<td>*</td>
</tr>
<tr>
<td>Exercise</td>
<td>0.164</td>
<td>0.028</td>
<td>96</td>
<td>*</td>
</tr>
</tbody>
</table>

**Note.** All listed categorical effects were, as a group, reliably different from zero. A superscript 1, 2, or 3 indicates that the effect size was statistically greater (after Bonferroni correction) than the effect size for the 1st, 2nd, or 3rd item, respectively, listed in that category (e.g., a “1,3” superscript means that the value in that cell was statistically greater than the 1st and 3rd listed items in that category). Asterisks indicate which categories were significantly different from zero.

---

**Training characteristics**

Interestingly, participants in combined strength and aerobic training regimens improved to a reliably greater degree than those in aerobic training alone ($0.59 vs. 0.41, SE = 0.043, n = 101, p < .05$). Also, participation in relatively brief training programs provided at least as much benefit as moderate training, but not quite as much as long-term training programs. Finally, short bouts of exercise (<30 min) had very little impact on cognitive function; the effect at this training duration was not significantly different from zero (see Table 1).
adults than younger adults in the regions of brain that are recruited to carry out a variety of cognitive tasks (Cabeza, 2001). One suggested explanation for this finding is that older adults recruit additional cortical areas to compensate for losses in neural efficiency. Another view characterizes dedifferentiation as a simple marker of cognitive decline. Some evidence in support of the compensatory hypothesis has been provided by studies that have examined the relationship between performance and brain activation. For example, Rympa and D'Esposito (2000) found, in an event-related fMRI study, that higher levels of activation of dorsolateral prefrontal cortex were associated with faster working memory retrieval for older adults.

Longitudinal assessments of cardiovascular changes and neurocognitive functioning would allow one to test the role that dedifferentiation plays in normal aging more directly. Such assessments would enable researchers to determine whether improvements in cognitive function that result from enhanced cardiovascular fitness would lead older adults to become more dissimilar from younger adults in their patterns of brain activation (i.e., increased dedifferentiation). Alternatively, cardiovascular improvements might “turn back the clock,” biologically speaking, and lead to patterns of neural activation that are more similar to the pattern of young adults.

The finding of significant effects for programmatic and demographic moderators also provides important information concerning potential boundary conditions on the fitness-cognition relationship, and suggests additional questions for further research. For example, it will be important to determine whether the larger fitness benefit for older than for younger senior citizens is the result of age differences in general health or education, or is instead a function of baseline cognitive and fitness levels. Similarly, the neuroprotective role of estrogen (Garcia-Segura, Cardona-Gomez, Chowen, & Azcoitia, 2000) and estrogen replacement therapy is an important topic for further research, in light of the fact that the fitness-related cognitive benefits were larger for women than for men. Also, the results reported here suggest that even clinical populations of older adults can benefit cognitively from physical exercise. Unfortunately, the relatively small number of published clinical studies prevents closer examination of the moderating effects of individual physical or cognitive maladies on the efficacy of the training programs. Further research into this issue is clearly important and much needed. The findings regarding the moderating effects of the type of fitness training, program duration, and training-session duration indicate that these factors should be systematically examined in future intervention studies.

**REFERENCES**


---

**Fig. 1.** Effect sizes for the different process-task types reflecting the four theoretical hypotheses concerning the process-based specificity of the benefits of fitness training. Parenthetical notations on the x-axis indicate the number of effect sizes contributing to the point estimates for each task type in the exercise (E) and nonexercise (C) groups. Error bars show standard errors.
Fitness and Cognition


*Study included in the meta-analysis.*

(Received 11/8/01; Revision accepted 4/8/02)