First publ. in: The Journals of Gerontology 65A (2010), 6, pp. 680-688

doi:10.1093/gerona/glq053

Complex Mental and Physical Activity in Older Women and Cognitive Performance: A 6-month Randomized Controlled Trial

Verena Klusmann,[1] Andrea Evers,[1] Ralf Schwarzer,[2] Peter Schlattmann,[3] Friedel M. Reischies,[1] Isabella Heuser,[1] and Fernando C. Dimeo[4]

[1]Department of Psychiatry, Charité Universitätsmedizin CBF, Berlin, Germany.
[2]Department of Health Psychology, Freie Universität Berlin, Germany.
[3]Department of Biometrics, Charité Universitätsmedizin CCM, Berlin, Germany.
[4]Section of Sports Medicine, Charité Universitätsmedizin CBF, Berlin, Germany.

Address correspondence to Verena Klusmann, MSc, Department of Psychiatry, Charité Universitätsmedizin Berlin CBF, Eschenallee 3, D-14050 Berlin, Germany. Email: verena.klusmann@charite.de

Background. Several reports suggest beneficial impacts of either physical or mental activity on cognitive function in old age. However, the differential effects of complex mental and physical activities on cognitive performance in humans remain to be clarified.

Methods. This randomized controlled trial evaluates a cognitive and a physical standardized 6-month activity intervention (3 × 1.5 h/wk) conducted in Berlin (Germany). Two hundred fifty nine healthy women aged 70–93 years were randomized to a computer course (n = 92), an exercise course (n = 91), or a control group (n = 76), of whom 230 completed the 6-month assessment. Group differences in change over a period of 6 months in episodic memory (story recall, possible range, 0–21; word recall, possible range, 0–16), executive control (working memory, ie, time quotient of Trail Making Tests B/A), and verbal fluency were evaluated by analyses of covariance (intention to treat) adjusting for baseline, fluid intelligence, and educational level.

Results. In contrast to the control group, both the exercise group, ΔM (SD) = 2.09 (2.66), $p < .001$, and the computer group, ΔM (SD) =1.89 (2.88), $p < .001$, showed improved delayed story recall. They maintained performance in delayed word recall and working memory (time measure) as opposed to the control group that showed a decline, ΔM (SD) = −0.91 (2.15), $p = .001$, and ΔM (SD) = 0.24 (0.68), $p = .04$, respectively.

Conclusions. In healthy older women, participation in new stimulating activities contributes to cognitive fitness and might delay cognitive decline. Exercise and computer classes seem to generate equivalent beneficial effects.

Key Words: Mental activity—Physical activity—Cognitive performance—Older adults.

LIFE of the elderly is characterized by a high risk of multimorbidity accompanied by a high prevalence of cognitive impairment short of dementia. A decline in fluid cognitive abilities, such as perception, reasoning, and memory, is part of the normal cognitive aging process (1). There is mounting evidence from epidemiological studies, experimental trials, and basic neurosciences that this age-related cognitive decline can be positively offset by intellectually stimulating activities (2). To date, most intervention studies have focused on either the effects of specific mental training or the benefits from physical exercise. It was found that cognitive training is especially effective when particular cognitive skills are practiced (3,4). However, findings concerning the effects of physical activity on cognitive performance have been mixed. Whereas some studies reported preventive effects of physical activity, others were inconclusive (5,6). Meta-analyses have found a moderate overall positive effect from physical training on cognition, with the largest effects occurring in motor function and auditory attention, followed by cognitive speed and visual attention (7–9). Data on the relationship between cardiovascular fitness benefits and cognitive improvements were inconclusive (9,10).

Animal research provides support for a positive effect of physical activity on brain function (11). Physical activity has been shown to have an acute upregulating effect on neurogenesis and to result in higher concentrations of brain-derived neurotrophic factor (12,13). The survival of newborn cells in the hippocampus was elevated in old mice being placed in an enriched environment, leading to significantly better cognitive performance in comparison to standard keeping (14).

Very few studies in humans have compared the effects of cognitive and physical stimulation. Oswald et al. (15) found that cognitive training alone, but not physical training alone, led to transiently improved cognitive function, whereas the combination of cognitive with physical training translated to longer lasting improvement. Similarly, in a small sample, Fabre et al. (16) found evidence pointing to an advantage of combined aerobic and mental training over using either technique alone in a memory quotient. Neither of these studies has focused on relatively complex cognitive stimulation based on the enriched environment theory. Also, there is a general quest for "modernized" cognitive interventions that support goals of lifelong learning by navigating new technologies in cross-modal interventions (17). Concerning physical activity, multifaceted interventions that integrate strength and flexibility into aerobic fitness trainings were found to have greater positive effects on cognition (18), and the proposal for sufficiently broadly based health interventions was given support (19).

Thus, in our project, 259 healthy elderly women were encouraged to engage in new activities of either a physical or a mental nature, that is, an exercise or a computer course. Participants were randomized to either one of two group interventions or to a control group of 6 months duration. It was hypothesized that both intervention groups would benefit from the new challenging activity with regard to cognition, thus showing a favorable development of cognitive performance over a period of 6 months in contrast to the controls.

Methods

Participants

This study enrolled German-speaking women from Berlin, who were older than 70 years. Eligibility criteria were (a) being unfamiliar with the computer and (b) exercising less than 1 h/wk. Criteria for exclusion included severe visual or hearing impairment or a previous or current diagnosis of depression or psychosis, or any other neurological or medical disorder that would interfere with cognitive performance or preclude successful participation in the intervention programs. Participants were screened to rule out the presence of cognitive impairment or depression and were included if they made no more than four errors on the Mini-Mental State Examination (MMSE; 20) and scored less than six points on the 15-item short-form Geriatric Depression Scale (GDS-SF; 21). Eligible women underwent the interventions consecutively in seven cohorts of about 34 persons each. Written informed consent was obtained, and the local ethical review board had approved the study protocol.

Interventions

One intervention comprised physical exercise and the other a complex cognitive task, whereas the control group got the instruction to live their habitual life. Since the world of computers is mainly an untapped area in the group of older people, especially in women, a computer course seemed particularly suitable for a new cognitive, cross-modal challenge.

The exercise program consisted of aerobic endurance, strength, and flexibility training, as well as practice of balance and coordination. Typically, exercise sessions started with 30 minutes endurance training on bicycle ergometers or treadmills with pulse monitors. The computer course covered heterogeneous and multifaceted themes including creative matters as well as coordinative and memory tasks, for example, learning how to operate with the common software and hardware, writing, playing, calculating, surfing on the Internet, e-mailing, drawing, image editing, and video taping.

The courses were carried out in different districts of Berlin. Public buildings, such as schools and fitness centers, served as study sites, where groups of twelve women were instructed. For both interventions, standardized manuals were developed by a certified sports physician and by an experienced computer teacher for seniors. Both manuals entailed 75 intervention units of 90 minutes. The last few sessions covered additional materials allowing the manuals being fitted to each 6-month interval of the seven cohorts. Course instructors were carefully trained in the manual application, and they daily documented course attendance and reasons for absence.

Objectives

Reviewing previous research, we identified a lack of long-term experimental studies, especially a deficit in randomized controlled trials, which would allow for a conclusive statement about the causal relationship between engagement in new activities and cognitive abilities. In addition, there is paucity with regard to studies examining the analogous or diverse effects of cognitive vs physical activity. Hence, we directly compared a mentally and a physically active group with a control group.

It was assumed that both activities lead to favorable cognitive performance when compared with the control group. "Favorable" cognitive performance was defined as either (a) an improvement of performance in the intervention groups from pre- to post-testing or (b) maintenance of preintervention performance over a period of 6 months.

Clinical Evaluation

A first screening was conducted by phone. If provisionally eligible, participants were invited for a medical evaluation. During a 1-hour session, a medical history as well as a 6-minute walking test (assessing meters completed) and a resting electrocardiogram (ECG) were done. Blood was drawn for the measurement of routine laboratory parameters to rule out chronic disorders (ie, chronic renal failure,

Table 1. Baseline Neuropsychological Assessment

Measure	Objectives
Consortium to Establish a Registry for Alzheimer's Disease (CERAD; 22)	General geriatric cognitive status
Mini-Mental State Examination (20)	Dementia screening
Naming semantic category members	Semantic verbal fluency
Boston Naming Test	Visual confrontation naming
Figures drawing	Constructional praxis
Word list recall	Episodic memory
Lector Test (23)	Educational level
LPS-3/50+ (24)	Fluid intelligence
Rivermead Behavioural Memory Test (25), subtest: story recall	Episodic memory in a naturalistic task
Free and Cued Selective Reminding Test (26)	Episodic memory in a classical experimental condition
Reitan Trail Making Tests A&B (27)	Speed and executive function: working memory
Stroop Test (28)	Executive attention: inhibition of irrelevant information

Note: LPS-3/50+ = Leistungs-Prüf-System (Performance Test System).

hypothyroidism). Women meeting the eligibility criteria for study participation were randomized to one of the intervention groups or the control group. Those being randomized to the exercise group had to pass an additional stress ECG to determine the individual optimal training heart rate. After 6 months of participation, the second clinical evaluation was done, including the 6-minute walking test, blood drawings, and a resting ECG.

Neuropsychological Evaluation

Before participants were informed about their group assignment, at a second 2.5-hour appointment, an extensive battery of neuropsychological tests (Table 1) was administered by a neuropsychologist who was blinded to randomization and group membership.

The neuropsychological assessment had a standardized format and was conducted in a face-to-face testing situation applying paper-and-pencil tests. The general geriatric cognitive status (22), as well as the educational level (23) and fluid intelligence (24) were screened at baseline. The verbal fluency assessment (22), the Rivermead Behavioural Memory Test (RBMT; 25), the Free and Cued Selective Reminding Test (FCSRT; 26), the Trail Making Tests (TMT A&B; 27), and the Stroop Test (28) were of particular interest. As primary outcomes these were applied at baseline and repeated at 6-month follow-up, assessing episodic memory, working memory, and executive attention, respectively. For this purpose, parallel versions for RBMT, FCSRT, TMT A&B, and verbal fluency (animals, food) were administered in random order. Also, demographic characteristics and self-reported exercise levels (defined as hours per month spent with sports implying quickened pulse and perspiration) were recorded. All assessors had undergone intensive training for test application including coding procedures.

Sample Size

Sample size was calculated using G*Power 3 (http://www.psycho.uni-duesseldorf.de/aap/projects/gpower/), assuming a half standard deviation of FCSRT scores in the 70–79-year-olds, M (SD) = 31.33 (6.74), of the Berlin Aging Study as a representative, meaningful intervention effect (29). Fixing α at 5% and power at 80%, $n = 177 = 3 \times 59$ was calculated as the necessary sample size for a comparison of three groups, half a $SD = 3.37, f = 0.24$; $F_{crit}(2,174) = 3.05$, $\lambda = 9.83$. A dropout rate of approximately 25 participants per group was expected, which led to a final sample size calculation of $n = 3 \times 84 = 252$. Actually, 259 female volunteers were recruited.

Randomization

The randomization sequence for each of the seven study cohorts was generated using Research Randomizer (www.randomizer.org) by VK. A study assistant prepared seven sets of 34 numbered envelopes containing the accordant randomization results (12 for the intervention groups each and 10 for the control group). We chose to include slightly more participants into the intervention groups because we expected a higher dropout rate in these groups than in the control group.

The sealed envelopes, all prepared before starting the examination of the first cohort and kept locked in a safe-deposit box, were given on a daily basis to the study nurse in consecutive order. Envelopes were opened after the main part of the clinical baseline evaluation to have the participants of the exercise group undergo the additional stress ECG. If study candidates withdrew from the study or were excluded because of lacking eligibility criteria at a later point in time, the study assistant prepared additional envelopes containing the corresponding assignments of those who dropped out in the sequence of deposit. Participants and neuropsychological assessors were blinded to group allocation up to the completed baseline examination of the whole cohort (double blind); participants were then informed by mail. Assessors were kept blind at post-test by explicitly instructing the participants not to discuss any of the information regarding randomization and intervention with the research staff conducting the testing. Staff members consigned with the scheduling of participants were not involved in the randomization procedure. We are not aware of any breaches of protocol.

Statistical Analyses

For all tests used to assess differences in cognitive performance from baseline to follow-up, change scores were calculated, that is, the difference between baseline and follow-up measurements (post-test minus pre-test scores). Arithmetic means and standard deviations were computed for change scores because these were normally distributed continuous variables. Analyses of covariance (ANCOVA) with planned contrasts on change scores with 95% confidence intervals (CI) were applied. Baseline scores of the pre–post measures and those obtained for educational level (Lector Test; 23) and general fluid intelligence (LPS-3/50+, Leistungs-Prüf-System/Performance Test System; 24) were included as covariates as preset prior to data collection. Treatment effects were obtained by *B*-estimates being provided by planned contrasts of ANCOVA. To evaluate change over the 6-month period within the groups, in a second step, dependent *t*-tests on pre- and post-scores were calculated. Strengths of associations were documented by the effect size partial η^2, calculated as $\eta^2 = SS_{effect}/(SS_{effect} + SS_{error})$, SS being the sum of squares. Following the convention of Cohen (1988), $\eta^2 \geq .01$ is regarded as small effect, $\eta^2 \geq .06$ as medium effect, and $\eta^2 \geq .14$ as large effect (30). For the meters completed in the 6-minute walking, test change scores (post-test minus pre-test meters) were calculated as for the cognitive performance measures. Analyses of variance (ANOVA) with planned contrasts on these change scores with 95% CI were conducted. To address possible problems of multiple comparisons by an inflated type I error rate, we applied the Bonferroni procedure to test conservatively for the global null hypothesis at $\alpha = 0.05$ for the between-group differences in cognitive performance change. The overall null hypothesis implies that there is no group difference at all and it can be rejected if at least one test is significant at adjusted $\alpha' = \alpha/m$, where m is the number of comparisons conducted.

Results

Flow of Participants

Fifty-four of 313 enrolled women did not meet the eligibility criteria, leaving 259 participants to be randomized (91 for the exercise, 92 for the computer, and 76 for the control condition), of whom 12 participants (5 of the exercise and 7 of the computer condition) refused to participate after being informed about their group assignment and withdrew consent before treatment started. Thus, 247 (95.4% of randomized participants; ie, 86 for the exercise, 85 for the computer, and 76 for the control condition) women were allocated to the corresponding groups, of whom 230 (93.1% of baseline, 88.8% of randomized) returned for follow-up (Figure 1). Women refusing to participate in the allocated intervention program and women lost to follow-up did not significantly differ from the remaining sample regarding sociodemographic variables, general cognitive status, and baseline scores of neuropsychological measures. There was one exception in the control group only: Women lost to follow-up performed significantly worse compared with those control women available to follow-up in free recall long delay. Models were fitted to available data, not imputing missing data. Thus, 80 women in the exercise intervention group, 81 women in the computer course group, and 69 women in the control group were available for the intention-to-treat analyses, including the women who discontinued the intervention at any point in time but were available to follow-up. Among the reasons for discontinuations were overcommitment, acute diseases, and death of one woman. Three women of the computer group were excluded from analyses of pre–post change in one cognitive test each, due to incorrect test data assessment. Four women of the computer condition, three women of the exercise condition, and three women of the control group were excluded from the analyses of pre–post change in the 6-minute walking test due to missing data.

Dates of Recruitment and Follow-up

The baseline evaluation period of a cohort lasted approximately 4 weeks and began 5 weeks before treatment allocation. The actual group interventions started 1 week later (mean period between baseline and intervention start 29.6 days; median 27 days). After termination of the 6-month group interventions, follow-up evaluations were carried out mostly within the subsequent 2 weeks (mean follow-up 10.9 days; median follow-up 7 days). Successive recruitment commenced in June 2006 and the first intervention groups in September 2006. Follow-up measures of the seventh cohort started in April 2008 and were completed by the end of that month.

Characteristics of the Study Population

Participants showed no signs of cognitive impairment (MMSE *M* [*SD*] = 28.78 [0.96], Min = 26) or depression (GDS-SF *M* [*SD*] = 1.83 [0.43], Max = 5); none of the women received antidementia or antidepressant drugs. There were no apparent group differences at baseline regarding age (range was 70–93 years), demographic characteristics, general cognitive status, and performance in any of the applied cognitive measures (Table 2), as well as self-reported exercise levels and the meters completed in the 6-minute walking test.

Intervention Effects: Group Differences in Change Scores in Cognitive Performance

ANCOVA with planned simple contrasts on the change scores in cognitive test performance from baseline to follow-up with 95% CI revealed that both intervention groups

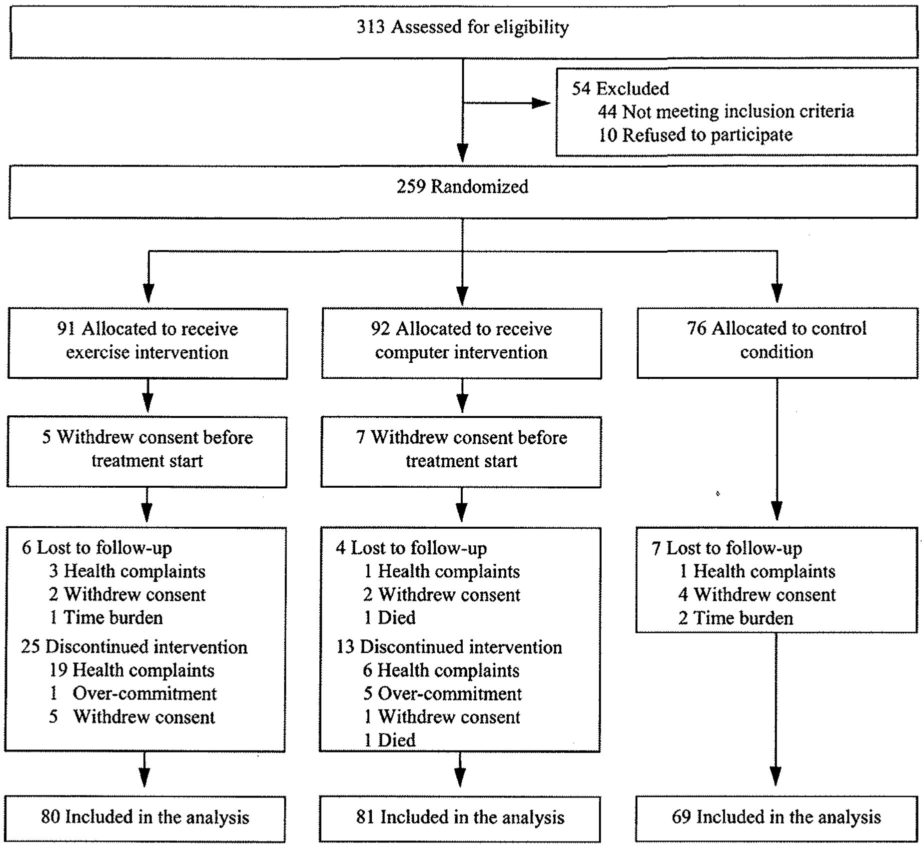

Figure 1. Flow of participants.

exhibited better development of performance than the control group in immediate and delayed recall in RBMT, $F(2,224) = 5.14$, $p = .007$, $\eta^2 = .04$ and $F(2,224) = 4.44$, $p = .01$, $\eta^2 = .04$, respectively. This was also true for performance in free recall FCSRT long delay, $F(2,224) = 4.14$, $p = .02$, $\eta^2 = .04$, and in working memory in TMT B/A, $F(2,223) = 3.31$, $p = .04$, $\eta^2 = .03$. There were no significant group differences in change scores regarding performance in free recall FCSRT short delay, $F(2,224) = 2.91$, $p = .06$, $\eta^2 = .03$, Stroop Test, $F(2,222) = 0.78$, $p = .46$, and semantic verbal fluency, $F(2,224) = 0.55$, $p = .58$. Bonferroni procedure to test the global null hypothesis revealed that with $\alpha = 5\%$ and seven cognitive tests, $m = 7$, we get $\alpha' = \alpha/m = 0.05/7 = 0.007$. Thus, we can reject the global null hypothesis with at least one test being significant at $p = .007$, namely the immediate recall in RBMT. Effects were of small size ($.01 < \eta^2 < .06$; see Table 3).

In addition, change scores in the 6-minute walking test were significantly different in the three groups, $F(2,217) = 3.75$, $p = .03$, $\eta^2 = .03$: the exercise group ($n = 77$; change ΔM [SD] = 51.30 [75.87]) outperformed both, the controls ($n = 66$; ΔM [SD] = 20.26 [89.27]; contrast to exercise group: $B = -31.04$, $SE = 13.28$, 95% CI, −57.21 to −4.87, $\eta^2 = .03$), and the computer group ($n = 77$; ΔM [SD] = 21.03 [72.93]; contrast to exercise group: $B = -30.27$, $SE = 12.76$, 95% CI, −55.42 to −5.13, $\eta^2 = .03$).

The evaluation of change over the 6-month period within the groups applying dependent t-tests on pre- and post-scores (Table 3) yields that both intervention groups showed an increase in immediate and delayed story recall (RBMT) from pre- to post-testing, whereas the control group, in contrast, had similar pre- and post-test scores. The increase in both intervention groups in story recall was large ($\eta^2 > .14$). In practical terms, the increase of the exercise group in immediate and delayed story recall was approximately 26% and 40%, respectively, whereas for the computer group it was roughly 30% and 35%. Regarding FCSRT long delay, the gains for the intervention groups in contrast to the control group were due to a significant decrease of performance in the control group from baseline to 6-month follow-up (approximately 6.5%), whereas both intervention groups maintained their preintervention levels. Similarly, the advantage of the intervention groups in performance on TMT B/A resulted from a significant decrease in performance in the control group (approximately 10%), while the computer course maintained the preintervention performance level and the exercise group actually increased slightly (approximately 7%), but this effect did not reach statistical significance. Decreases

Table 2. Characteristics of the Study Population

Demographics	Total sample (N = 259)	Exercise group (n = 91)	Computer group (n = 92)	Control group (n = 76)
Age (years)	73.6 (4.2)	73.6 (4.0)	73.6 (4.4)	73.5 (4.3)
Marital status				
Married (%)	57 (22.0)	24 (26.4)	15 (16.3)	18 (23.7)
Widowed (%)	95 (36.7)	32 (35.2)	33 (35.9)	30 (39.5)
Divorced (%)	71 (27.4)	25 (27.5)	27 (29.3)	19 (25.0)
Other (%)	36 (13.9)	10 (10.9)	17 (18.5)	9 (11.8)
Education (years)	12.0 (2.6)	11.8 (2.5)	12.0 (2.6)	12.0 (2.8)
Fluid intelligence (LPS-3/50+)	18.82 (4.32)	18.64 (3.91)	19.10 (4.57)	18.70 (4.51)
Educational level (Lector Test)	40.45 (4.18)	40.31 (4.01)	40.50 (4.86)	40.57 (3.47)
Neuropsychological tests				
RBMT immediate story recall	6.56 (2.34)	6.27 (2.30)	6.45 (2.20)	7.04 (2.50)
RBMT delayed story recall	5.49 (2.12)	5.17 (1.88)	5.50 (2.14)	5.87 (2.33)
Free word recall FCSRT short delay	35.15 (4.69)	35.43 (4.03)	34.89 (4.69)	35.12 (5.41)
Free word recall FCSRT long delay	13.68 (1.81)	13.46 (1.60)	13.73 (1.62)	13.87 (2.22)
Semantic verbal fluency	24.68 (5.00)	24.85 (5.53)	24.29 (4.97)	24.93 (4.37)
Stroop Test	25.46 (10.66)	25.86 (11.72)	24.43 (9.42)	26.22 (10.79)
TMT B/A	2.45 (0.84)	2.57 (0.78)	2.40 (0.98)	2.36 (0.70)
General cognitive status (CERAD)				
Boston Naming Test	14.20 (1.02)	14.09 (1.08)	14.20 (1.01)	14.34 (0.93)
Mini-Mental States Examination	28.76 (0.97)	28.80 (0.89)	28.84 (0.94)	28.62 (1.08)
Figure items drawn	10.18 (1.22)	10.18 (1.23)	10.18 (1.21)	10.18 (1.24)
Figure items drawn delayed	8.60 (2.13)	8.41 (2.14)	8.66 (2.21)	8.75 (2.03)
Total words recalled immediately	21.17 (3.07)	21.01 (3.09)	21.18 (3.34)	21.34 (2.71)
Total words recalled delayed	7.04 (1.76)	7.08 (1.75)	6.90 (1.83)	7.16 (1.72)
Physical exercise and fitness				
Self-reported exercise level	1.32 (1.83)	1.50 (1.91)	1.36 (1.81)	1.08 (1.73)
Meters completed in 6-min walking test	354.20 (107.10)*	354.37 (106.90)*	351.89 (108.61)	356.80 (106.86)

Notes: Values are means (SD) unless stated otherwise. FCSRT = Free and Cued Selective Reminding Test; LPS-3/50+ = Leistungs-Prüf-System (Performance Test System), Version 50+; RBMT = Rivermead Behavioural Memory Test; TMT = Trail Making Tests. For most measures, higher scores indicate better performance. The range of scores on the LPS-3/50+ is 0–40; the range of scores on the Lector Test is 0–48; the range of scores on RBMT story recall is 0–21; the range of scores on the free recall FCSRT short delay is 0–48 (sum score of three recall trials short delay separated by distractors) and on the free recall FCSRT long delay is 0–16; the range of scores on fluency is 0 to open end; the range of scores on the Boston Naming Test is 0–15; the range of scores on Mini-Mental State Examination is 0–30, with scores from 30 to 26 indicating no cognitive impairment; the range of scores on figure items drawn is 0–11; the range of scores on total words recalled immediately is 0–30 (sum score of three consecutive immediate recall trials) and on total words recalled delayed is 0–10. For the following measures, higher scores indicate worse performance: The Stroop Test score is measured by time for task performance in the color–word interference task less the time needed to perform the simple color task; TMT B/A scores are a quotient of time needed to solve part B over time needed to solve part A. Self-reported exercise levels were defined as hours per month spent on sports implying quickened pulse and perspiration.

*Total n = 258, exercise group n = 90, due to missing data.

in FCSRT long delay and working memory (TMT B/A) in controls were of medium ($\eta^2 > .06$) to large size ($\eta^2 > .14$). In all three groups, FCSRT short delay and semantic verbal fluency remained stable, while the Stroop task performance improved significantly (small to medium effects; see Table 3).

Discussion

Using either a stimulating physical or mental activity, this 6-month intervention study demonstrates that in healthy, older women, naturalistic episodic memory function increased while episodic memory in the classical experimental condition and working memory were maintained. Interestingly, the exercise group and the computer course group were not found to differ in any of the studied cognitive abilities. The question whether the effects of the two training types would be additive or interactive has not been addressed. Finally, no statistically significant differences in change in semantic verbal fluency and in suppression of irrelevant information were found between the three groups.

This direct head-to-head comparison between a physical and a mental intervention demonstrates that participation in either of both new activities results in gains in, or maintenance of, cognitive performance in older women. The common denominator of these activities was the management of new complex situations. From preclinical studies in laboratory animals it is well known that an enriched environment has neuroprotective effects. Appropriate intellectual stimuli and specific cognitive effort are necessary to make use of this potential (31). Physical or cognitive activities in the present study might tap this potential, thus improving the plasticity of the aging brain to reduce cognitive senescence in humans.

One explanation of the gains in naturalistic episodic memory of the intervention groups is that both activity types augment strategy learning for everyday memory because the task is quite similar to daily life activities (32). Regarding

Table 3. Activity Intervention Effects: Comparison of Change in Cognitive Performance

ANCOVA Main Result	Group	M (SD) Pre	M (SD) Post	Change ΔM (SD)	Between-Group Contrast (ANCOVA), Intervention vs. Control Group: B (95% CI)	SE	η^2	Within-Group Contrast (t-tests), Baseline to Follow-up: t	df	p	η^2
RBMT immediate recall; $F(2,224) = 5.14$,	Exercise group ($n = 80$)	6.38 (2.35)	8.06 (2.60)	1.69 (2.57)	1.02 (0.23 to 1.81)	0.40	.03	5.86	79	<.001	.30
$p = .007$, $\eta^2 = .04$	Computer group ($n = 81$)	6.32 (2.21)	8.22 (2.94)	1.90 (2.93)	1.22 (0.43 to 2.01)	0.40	.04	5.84	80	<.001	.30
	Control group ($n = 69$)	7.12 (2.49)	7.32 (2.28)	0.20 (3.02)				0.56	68	.58	
RBMT delayed recall;	Exercise group ($n = 80$)	5.26 (1.93)	7.35 (2.68)	2.09 (2.66)	1.11 (0.31 to 1.92)	0.41	.03	7.03	79	<.001	.38
$F(2,224) = 4.44$,	Computer group ($n = 81$)	5.38 (2.14)	7.27 (2.68)	1.89 (2.88)	1.02 (0.21 to 1.82)	0.41	.03	5.90	80	<.001	.30
$p = .01$, $\eta^2 = .04$	Control group ($n = 69$)	5.92 (2.32)	6.44 (2.30)	0.51 (3.10)				1.39	68	.17	
FCSRT short delay;	Exercise group ($n = 80$)	35.65 (3.89)	36.05 (4.25)	0.40 (3.43)				1.04	79	.30	
$F(2,224) = 2.91$,	Computer group ($n = 81$)	35.15 (4.63)	34.69 (5.21)	−0.46 (4.25)				−0.97	80	.34	
$p = .06$, $\eta^2 = .03$	Control group ($n = 69$)	35.48 (5.07)	34.52 (4.32)	−0.96 (5.01)				−1.59	68	.12	
FCSRT long delay;	Exercise group ($n = 80$)	13.48 (1.56)	13.65 (1.58)	0.18 (1.90)	0.67 (0.16 to 1.17)	0.26	.03	0.82	79	.41	
$F(2,224) = 4.14$,	Computer group ($n = 81$)	13.68 (1.64)	13.67 (1.61)	−0.01 (1.76)	0.62 (0.12 to 1.12)	0.25	.03	−0.06	80	.95	
$p = .02$, $\eta^2 = .04$	Control group ($n = 69$)	14.07 (1.87)	13.16 (1.70)	−0.91 (2.15)				−3.53	68	.001	.15
Semantic verbal fluency;	Exercise group ($n = 80$)	24.66 (5.63)	25.60 (5.34)	0.94 (6.00)				1.40	79	.17	
$F(2,224) = 0.55$,	Computer group ($n = 81$)	24.33 (5.02)	24.96 (4.92)	0.63 (4.37)				1.30	80	.20	
$p = .58$	Control group ($n = 69$)	25.10 (4.10)	25.07 (4.29)	−0.03 (4.70)				−1.16	68	.96	
Stroop Test;	Exercise group ($n = 80$)	25.16 (9.94)	23.50 (8.89)	−1.66 (7.17)				−2.07	79	.04	.05
$F(2,222) = 0.78$,	Computer group ($n = 79$)*	24.29 (9.38)	22.09 (8.63)	−2.20 (6.67)				−2.93	78	.004	.10
$p = .46$	Control group ($n = 69$)	26.23 (10.92)	23.09 (6.97)	−3.14 (8.11)				−3.22	68	.002	.13
Trail Making Tests B/A;	Exercise group ($n = 80$)	2.54 (0.75)	2.36 (0.67)	−0.18 (0.91)	−0.27 (−0.50 to −0.04)	0.12	.02	−1.76	79	.08	
$F(2,223) = 3.31$,	Computer group ($n = 80$)*	2.31 (0.71)	2.34 (0.78)	0.03 (0.90)	−0.25 (−0.47 to −0.02)	0.12	.02	0.29	79	.77	
$p = .04$, $\eta^2 = .03$	Control group ($n = 69$)	2.36 (0.68)	2.60 (0.83)	0.24 (0.68)				2.10	68	.04	.06

Notes: ANCOVA = analysis of covariance; *CI* = confidence interval of *B*; *df* = degrees of freedom; FCSRT = Free and Cued Selective Reminding Test; LPS-3/50+ = Leistungs-Prüf-System (Performance Test System), Version 50+; RBMT = Rivermead Behavioural Memory Test; *SD* = standard deviation; *SE* = standard error of *B*. Results of ANCOVA with simple contrasts. Covariates included educational level (Lector Test), fluid intelligence (LPS-3/50+), and corresponding baseline measure. Planned simple between-group contrasts were reported only if the main effect of group on a measure was significant. *B* signifies a benefit over the control group in the scale unit of a corresponding measure (unstandardized parameter estimate) adjusted for covariates. Δ*M* denotes mean pre–post change, that is, mean of post-test minus pre-test scores, η^2 denotes the effect size ($\eta^2 = .01$, small effect; $\eta^2 = .06$, medium effect; $\eta^2 = .14$, large effect). Listed between-group contrasts were significant at $p < .05$.

*Three women of the computer group were excluded from analyses of pre-post change in one cognitive test each ($n = 2$ in Stroop Test; $n = 1$ in Trail Making Tests) due to incorrect test data assessment.

episodic memory in the classical experimental condition and working memory, maintenance of performance might have been achieved because activity and the inherent demand to face new challenges may have led to cognitive flexibility and an improved use of compensatory strategies in neuropsychological functions. Thus, we can conclude that learning results derived from the activities generalize to performance on specific laboratory cognitive tasks.

Results from this intervention study agree with findings from epidemiological prospective studies suggesting that physical or cognitive activity has preventive effects on age-related cognitive decline (33). By and large, our results are consistent with reviews concluding that cognitive maintenance and age-related decline can be positively influenced by intellectually stimulating activities and fitness training (2,9). However there are no meta-analyses available that report on broader cognitive challenges, let alone the simultaneous examination of physical and cognitive stimulation in the same analysis. The most exhaustive meta-analysis on the effects of physical activity examined supervised intervention studies in people older than 55 years and found that executive functions benefit the most (7). In addition to this, we demonstrated pronounced effects in two measures of episodic memory—a domain that was not considered as possible outcome in the meta-analysis by Colcombe and Kramer (7). Due to varying methodology, it is hard to draw firm conclusions from comparison with other studies, but possibly the long program duration and the intensity of our courses resulted in the distinct effects of our study.

In a recent intervention study, Lautenschlager et al. studied a sample of 170 memory complainers, most of whom were mildly cognitively impaired. They found that a 6-month physical activity program provided modest improvements in the cognitive subscale of the Alzheimer Disease Assessment Scale (5). Contrary to their study, in our trial, the implemented activities were indeed new for the participants and we compared a physically with a mentally challenging activity. Through the supervised group interventions that were stipulated in the manuals it was ensured that all participants got the same standardized treatment and adhered to the intervention. It is noteworthy that our sample consisted of healthy, well-functioning women without any signs of cognitive impairment or dementia. Despite this, the interventions showed potential to either increase memory or prevent its decline. Thus, our findings are promising in the context of risk reduction of dementia and are consistent with findings showing that not only cognitive training but also education and an engaged lifestyle might reduce the risk of dementia (34,35).

Dropout rates were relatively low and compliance quite high in our study (cf. Figure 1). This was most likely due to our procedure of contacting all participants, including those of the control group, on a regular basis by phone. Such brief but regular incentives should be implemented in health programs along with stringent documentation of attendance to endorse satisfactory compliance.

The fact that dropouts in the control group performed significantly worse in one of the measures of memory (free recall long delay) at baseline argues in favor of our interventions, strengthening the notion that our interventions actually led to an advantage over the control group. Those dropouts in the control group with low baseline performance likely would have increased the gap in change scores even more between the control and the experimental groups. Thus, results obtained for participants completing the study appear to be internally valid.

The performance on the Stroop task measuring executive attention, that is, organizing inhibition of irrelevant information, was not influenced by the interventions. The finding that all three groups increased in performance points to a practice or retest effect, which is most probably due to the fact that at post-test participants are already familiar with the task. This would put the usefulness of the Stroop test for the assessment of change into question, especially because a reasonable parallel version that could abolish the practice effect practically cannot be constructed due to the nature of the test.

Limitations

There are some limitations of our study that need to be addressed. First, we studied different cognitive outcomes in only one sample. However, it was our goal to apply two broad, rather unspecific and very different interventions and to evaluate their effects on cognitive domains known to be age sensitive. This issue might raise concerns about bias by inflated alpha errors due to multiple comparisons done within one sample. To address this possible problem, we confirmed the significance of our results by applying the Bonferroni procedure.

Second, one might be surprised that there was a decline in test scores of the control participants after only half a year. One might suspect that participants of the control group were less motivated to perform their best. However, there were no women that appeared to be less motivated at post-testing; on the contrary, it was impressive how competitive all women were. Thus, we reason that this finding is best explained by a small but significant aging-associated loss.

Third, no statistically significant differences have been detected between the two experimental groups. One might argue that the force driving equal benefits was the social contact both intervention groups enjoyed. We cannot refute this completely; however, prior intervention studies have revealed no differences in cognitive improvement between a social contact control group and a passive control group (36,37).

Finally, we used a single-, not a double-, blind design. However, to design a "placebo" control group would be methodologically challenging and, furthermore, to keep participants fully blinded would raise ethical questions.

Conclusions

Both activity programs in the present study are suitable for everyday life. Due to their nonspecificity, they provide the potential to be implemented in the daily routine of a large group of older people—an advantage over the specific cognitive training programs used in other studies. Moreover, engaging in a new and interesting cognitive or physical activity to maintain or improve cognitive fitness may not only lead to gains in cognitive test performance but may contribute to sustained independence. Fluid cognitive abilities, like episodic memory and executive function, have been linked to problem solving in daily life and maintenance of activities of daily living or improvements in everyday abilities (9,38,39). In this context it is noteworthy that especially the course teaching computer skills to novices was greatly appreciated by the older women who felt connected to modern technology. Biomedical models emphasize good mental functioning as an important determinant of successful aging while sociological models underline components such as life satisfaction, social functioning, and participation (40).

To our knowledge, the present study is the first to evaluate the effects of a multifaceted, new mental or physical activity in a large randomized controlled trial of older, well-functioning women. Taken together, our results suggest that different activity types, that is, mental activity or exercise, are equally suitable to support cognitive performance in older women. Challenge and novelty are presumed to be of particular importance. Becoming active in an exercise course or starting to work (and play) with the computer not only sustains cognitive performance but appears to help to postpone decline in those cognitive domains where pronounced age-related losses otherwise occur.

Funding

This work was supported by the German Research Foundation (grant number 429) Doctoral Program "Neuropsychiatry and Neuropsychology of Aging" (years covered: 2006–2008) to I.H. and the Gertrud and Hugo Adler Foundation (years covered: 2006–2008).

Acknowledgments

Senior coauthorship is shared by F.M.R., I.H., and F.C.D. I.H. is the head of the project group. The German Research Foundation and the Gertrud and Hugo Adler Foundation had no involvement in the study design; in the collection, analysis, and interpretation of data; and in the writing of this article and the decision to submit for publication.

References

1. Baltes PB, Staudinger UM, Lindenberger U. Lifespan psychology: theory and application to intellectual functioning. *Annu Rev of Psychol.* 1999;50:471–507.
2. Kramer AF, Bherer L, Colcombe SJ, Dong W, Greenough WT. Environmental influences on cognitive and brain plasticity during aging. *J Gerontol A Biol Sci Med Sci.* 2004;59:940–957.
3. Ball K, Berch DB, Helmers KF, et al. Effects of cognitive training interventions with older adults: a randomized controlled trial. *JAMA.* 2002;288:2271–2281.
4. Schaie KW, Willis SL. *Adult development and aging*, 5th ed. New York: Prentice-Hall; 2002.
5. Lautenschlager NT, Cox KL, Flicker L, et al. Effect of physical activity on cognitive function in older adults at risk for Alzheimer disease: a randomized trial. *JAMA.* 2008;300:1027–1037.
6. Churchill JD, Galvez R, Colcombe S, Swain RA, Kramer AF, Greenough WT. Exercise, experience and the aging brain. *Neurobiol Aging.* 2002;23:941–955.
7. Colcombe S, Kramer AF. Fitness effects on the cognitive function of older adults: a meta-analytic study. *Psychol Sci.* 2003;14:125–130.
8. Etnier JL, Salazar W, Landers DM, Petruzzello SJ, Han M, Nowell P. The influence of physical fitness and exercise upon cognitive functioning: a meta-analysis. *Sport Exerc Psychol.* 1997;19:249–277.
9. Angevaren M, Aufdemkampe G, Verhaar HJJ, Aleman A, Vanhees L. Physical activity and enhanced fitness to improve cognitive function in older people without known cognitive impairment. *Cochrane Database Syst Rev.* 2008;3:CD005381.
10. 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.
11. Hillman CH, Erickson KI, Kramer AF. Be smart, exercise your heart: exercise effects on brain and cognition. *Nat Rev Neurosci.* 2008;9:58–65.
12. Wiskott L, Rasch MJ, Kempermann G. A functional hypothesis for adult hippocampal neurogenesis: avoidance of catastrophic interference in the dentate gyrus. *Hippocampus.* 2006;16:329–343.
13. Cotman CW, Berchtold NC, Christie LA. Exercise builds brain health: key roles of growth factor cascades and inflammation. *Trends Neurosci.* 2007;30:464–472.
14. Kempermann G, Kuhn HG. More hippocampal neurons in adult mice living in an enriched environment. *Nature.* 1997;386:493–495.
15. Oswald WD, Gunzelmann T, Rupprecht R, Hagen B. Differential effects of single versus combined cognitive and physical training with older adults: the SimA study in a 5-year perspective. *Eur J Ageing.* 2006;3:179–192.
16. 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.
17. Acevedo A, Loewenstein DA. Nonpharmacological cognitive interventions in aging and dementia. *J Geriatr Psychiatry Neurol.* 2007;20: 239–249.
18. McAuley E, Kramer AF, Colcombe SJ. Cardiovascular fitness and neurocognitive function in older adults: a brief review. *Brain Behav Immun.* 2004;18:214–220.
19. Andrews GR. Care of older people: promoting health and function in an ageing population. *BMJ.* 2001;322:728–729.
20. Folstein MF, Folstein SE, McHugh PR. Mini-Mental State: a practical method for grading the state of patients for the clinician. *J Psychiatr Res.* 1975;12:189–198.
21. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clin Gerontol.* 1986;5: 165–173.
22. Morris JC, Heyman A, Mohs RC, Hughes JP. The consortium to establish a registry for Alzheimer's disease (CERAD): I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology.* 1989;39: 1159–1165.
23. Reischies FM, Wertenauer F, Kühl KP. Der "Lector" — ein Untersuchungsverfahren zur Bestimmung des verbalen Bildungsniveaus [The "Lector"—a testing method to determine verbal educational level]. *Nervenarzt.* 2005;76:843–855.
24. Sturm W, Willmes K, Horn W. *Leistungs-Prüf-System für 50-90jährige (LPS 50+): Testmappe [Performance test system for the 50 to 90-year-old (LPS 50+): testfolder].* Göttingen: Hogrefe; 1993.
25. Wilson BA, Cockburn J, Baddeley A. *The Rivermead Behavioural Memory Test.* Bury St. Edmunds, UK: Thames Valley Test Company; 1985.
26. Buschke H. Cued recall in amnesia. *J Clin Neuropsychol.* 1984;6: 433–440.
27. Reitan RM. Validity of the Trail Making Test as an indication of organic brain damage. *Percept Mot Skills.* 1958;8:271–276.
28. Oswald WD, Fleischmann UM. *Das Nürnberger Altersinventar (NAI) [The Nuremberg Aging Inventory].* Göttingen: Hogrefe; 1997.
29. Reischies FM, Lindenberger U. Grenzen und Potentiale kognitiver Leistungsfähigkeit im Alter [Limits and potentials of intellectual functioning in old age]. In: KU Mayer and PB Baltes (eds). *Die Berliner Altersstudie [The Berlin Aging Study].* Berlin: Akademie Verlag; 1996:351–377.
30. Cohen J. *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum; 1988.
31. Kempermann G. The neurogenic reserve hypothesis: what is adult hippocampal neurogenesis good for? *Trends Neurosci.* 2008;31: 163–169.
32. Bottiroli S, Cavallini E, Vecchi T. Long-term effects of memory training in the elderly: a longitudinal study. *Arch Gerontol Geriatr.* 2008;47: 277–289.
33. Wilson RS, Mendes De Leon CF, Barnes LL, et al. Participation in cognitively stimulating activities and risk of incident Alzheimer disease. *JAMA.* 2002;287:742–748.
34. Heyn P, Abreu BC, Ottenbacher KJ. The effects of exercise training on elderly persons with cognitive impairment and dementia: a meta-analysis. *Arch Phys Med Rehabil.* 2004;85:1694–1704.
35. Sitzer DI, Twamley EW, Jeste DV. Cognitive training in Alzheimer's disease: a meta-analysis of the literature. *Acta Psychiatr Scand.* 2006; 114:75–90.
36. Wadley VG, Benz RL, Ball KK, Roenker DL, Edwards JD, Vance DE. Development and evaluation of home-based speed-of-processing training for older adults. *Arch Phys Med Rehabil.* 2006;87:757–763.
37. Willis SL, Cornelius SW, Blow FC, Baltes PB. Training research in aging: attentional processes. *J Educ Psychol.* 1983;75:257–270.
38. Boron JB, Turiano NA, Willis SL, Schaie KW. Effects of cognitive training on change in accuracy in inductive reasoning ability. *J Gerontol B Psychol Sci Soc Sci.* 2007;62:179–186.
39. Ball K, Edwards JD, Ross LA. The impact of speed of processing training on cognitive and everyday functions. *J Gerontol B Psychol Sci Soc Sci.* 2007;62:19–31.
40. Bowling A, Dieppe P. What is successful ageing and who should define it? *BMJ.* 2005;331:1548–1551.