Improving cognition by adherence to physical or mental exercise: A moderated mediation analysis

Andrea Evers a; Verena Klusmann a; Ralf Schwarzer b; Isabella Heuser a
a Department of Psychiatry, Charité Universitätsmedizin Berlin, Campus Benjamin Franklin, Berlin, Germany
b Department of Health Psychology, Freie Universität Berlin, Berlin, Germany

Online publication date: 14 April 2011
Improving cognition by adherence to physical or mental exercise: A moderated mediation analysis

Andrea Evers*, Verena Klusmanna, Ralf Schwarzerb and Isabella Heusera

*Department of Psychiatry, Charité Universitätsmedizin Berlin, Campus Benjamin Franklin, Berlin, Germany; bDepartment of Health Psychology, Freie Universität Berlin, Berlin, Germany

(Received 11 June 2010; final version received 2 November 2010)

**Background:** The role of adherence to an intervention is examined to further understand the relationship between performing new challenging activities (either mental or physical ones) and their putative cognitive benefits.

**Method:** Healthy older women (N = 229, age range: 70–93 years) took part in a six-month randomised controlled trial, covering either a physical or mental activity (three × weekly). They completed five tests, measuring episodic and working memory pre- and post-intervention. A moderated mediation model was specified to test the strength of the indirect effect of the activity mode (i.e. physical vs. mental) through adherence (i.e. time spent on course attendance) on levels of baseline cognitive performance.

**Results:** Both physical and mental activity groups performed better over time than the control group (p < 0.001). Adherence predicted cognitive performance (p = 0.011). The indirect effect of the activity mode on cognitive performance through adherence was especially seen when levels of baseline composite scores were low (p = 0.023).

**Conclusion:** Older healthy women can improve episodic and working memory through spending time on a challenging physical or mental activity. Results are most promising for cognitively less fit women. Time spent on course attendance can be interpreted as an adherence indicator that makes a difference for various cognitive outcomes of the intervention.

**Keywords:** older women; cognitive improvement; adherence; spending time; intervention
baseline cognitive function and cognitive decline are negatively related; however, high intra-individual variation is described (Park, O’Connell, & Thomson, 2003). Behavioural brain reserve, which is also referred to as cognitive reserve (Scarmeas & Stern, 2003), could be increased by interventions to protect against this latter stated cognitive decline (Valenzuela & Sachdev, 2006).

In this article, we examine data of a previously reported RCTs to improve cognitive fitness (Klusmann et al., 2010). Healthy older women were randomised to a six-month standardised physical activity (i.e. an exercise course), mental activity (i.e. a computer course), or a passive control group. Participants in both intervention groups showed comparably better cognitive performance change over six months than the controls.

According to Hansen and McNeal (1996), who argue that interventions work by means of mediating variables, we explore the question of how the intervention was effective by assuming adherence to be the mediating variable. We further, address the question of whom this effect was especially seen by testing the moderating role of baseline cognitive performance. Thereby, we use a composite score of cognitive performance as the primary outcome measure. Figure 1 illustrates the three mechanisms, mediation, moderation and the integrated model of moderated mediation to be tested:

(i) testing the influence of the activity mode on cognitive change through adherence (Figure 1, Panel A, hypothesis on mediation),
(ii) examining the moderating role of cognitive baseline performance on the follow-up performance (Figure 1, Panel B, hypothesis on moderation),
(iii) testing the two assumptions simultaneously: (1) the activity mode has an indirect effect on cognition through adherence, and (2) the strength of the adherence–cognition association depends on the level of cognitive baseline performance (Figure 1, Panel C, hypothesis on moderated mediation). In other words, the hypothesised indirect effect (activity – adherence – cognitive outcome) would emerge at a certain level of baseline performance.

Methods

Participants and procedure

German-speaking women from Berlin, aged 70 years or older, were recruited by advertisements in newspapers, public transport systems and flyer distribution. Prior to randomisation, participants were screened to rule out a presence of cognitive impairment, depression or other neurological or medical diseases that would interfere with cognitive performance or course participation. Further, eligibility criteria were exercising less than one hour/week and being unfamiliar with a computer.

A total of 259 women aged 70 years and older were randomised to a six-month standardised physical activity (i.e. an exercise course, n = 86), mental activity (i.e. a computer course, n = 85), or control group (n = 76). Twelve women withdrew consent after being informed about group assignment, leaving 247 women who started the intervention. Both intervention programmes of physical or mental activities consisted of 90-minute standardised training sessions, three times weekly for six months. Participants were trained in groups of 12 by a certified exercise trainer or a qualified computer trainer in different locations throughout Berlin that were accessible by public transportation.

The exercise course started with 30 minutes of endurance training on bicycle ergometers or treadmills. Heart rate monitors were used to avoid exceeding the individually detected heart rate maximum. The endurance training was followed by 60 minutes of strength, flexibility, balance and coordination training. The computer course comprised learning how to operate with the common software and hardware (writing, playing, calculating, surfing the Internet, emailing, drawing, image editing and video taping).

On average, 73 training classes (range 70–74 sessions) were offered to a total of seven groups (physical and mental) that started interventions
successively. A total of 230 women, i.e. 80 women on the physical exercise course, 81 women on the computer course and 69 women in the control group, returned for follow-up assessments at the end of the 6 month training session. The data of this sample were used for all analyses in this article.

Cognitive assessment
As reported elsewhere (Klusmann et al., 2010), baseline and follow-up assessments covered neuropsychological tests measuring episodic memory, working memory, executive attention and semantic verbal fluency. Significant differences between the activity groups and the controls were found in five validated tests that were measuring episodic and working memory: the Free and Cued Selective Reminding Test (FCSRT; Buschke, 1984; Grober, Lipton, Hall, & Crystal, 2000), short- and long-delayed recall, measuring episodic memory in a classical laboratory condition, the Rivermead Behavioural Memory Test, immediate and delayed recall (Wilson, Cockburn, & Baddely, 1985), measuring episodic memory in a more naturalistic way, and the Reitan Trail Making Test (TMT; A & B; Reitan, 1955) measuring speed and executive function. All these tests represent well-known standard inventories in neuropsychological assessment.

Participants in our study were healthy and well-educated, and, therefore, ceiling effects are likely to play a role. To minimise ceiling artefacts and other sources of measurement errors and to test training effects at the level of one overall outcome rather than at the level of one single test, a composite score was calculated for women who completed all measures of those five tests at baseline (pre-score) and follow-up (post-score). Pre-scores of each single test were standardised (z-score) using the mean and standard deviation (SD) of the total sample and then averaged into one single score. Post-scores were standardised using the baseline mean and SD to facilitate interpretation of change. Averaging the standardised scores provided a composite score of cognitive performance for both measurement points with higher scores indicating higher cognitive performance. To assess differences in cognitive performance from baseline to follow-up in the cited tests, change scores were computed (raw post-scores minus raw pre-scores for each single test). These change scores were then converted to z-scores and these standardised scores were averaged into a composite measure of change of cognitive performance, with greater absolute values indicating greater change.

Adherence assessment
Course attendance in the activity classes was recorded by all trainers for each participant and intervention unit. Adherence was defined as the amount of time spent on participation in the courses plus the individual average travel time to course locations. Adherence (A) was calculated by multiplying course attendance (CA) with the sum of 90 minutes (i.e. duration of one intervention unit) plus travel time in minutes to and from course locations (TT):

\[ A = CA \times (90 \text{ minutes} + TT) \]

For reasons of convenience, the result was converted into units of hours. The average course attendance is the ratio of attended to offered course units in per cent.

Statistical analyses
Primary analysis on cognitive outcome
A 3 × 2 mixed ANCOVA was conducted for the cognitive composite score to test the main effect of all three experimental groups, the main effect of time and whether the change in composite scores was different for the three experimental groups (time × group interaction). Baseline measures for educational level (Lector Test; Reischies, Wertenauer, & Kühl, 2005) and general fluid intelligence (LPS-3/50; Leistungs-Prüf-System/Performance Test System; Sturm, Willmes, & Horn, 1993) were included as covariates as done in the primary analyses described in Klusmann et al. (2010). Additionally, dependent sample t-tests on composite scores from pre- to post-testing were conducted to examine changes within each experimental group.

(i) Hypothesis on mediation
Given a significant change on cognitive outcome in both intervention groups compared to controls, a simple mediation model (Figure 1, Panel A) was conducted by Sobel Z using an SPSS macro provided by Preacher and Hayes (2004) to examine the mechanism of the intervention through adherence. The term ‘mediation’ implies that a significant simple effect (\(c; X \rightarrow Y\)) is initially present (Holmbeck, 1997; Preacher & Hayes, 2004). If \(X\) and \(Y\) are not associated directly, they may be related indirectly through the intervening variable \(M\); this would be termed an indirect effect (Mathieu & Taylor, 2006).

The applied Sobel test is based on the assumption that this indirect effect (\(ab\) or \(c - c'\)) is normally distributed which is the case only in large samples (Preacher & Hayes, 2008). Thus, we additionally used a product-of-coefficients strategy with bootstrapping, a non-parametric method for assessing indirect effects, to test and estimate the indirect effect (Hayes, 2009; Preacher, Rucker, & Hayes, 2007; Shrout & Bolger, 2002).

(ii) Hypothesis on moderation
We predicted that the baseline composite score would moderate the relationship between adherence and outcome composite score (Figure 1, Panel B). We tested this moderator model in a multiple regression.
Therefore, a moderated mediation (termed alterna-
tively moderated indirect effect) was regressed on activity mode (physical vs. mental activity, $X$), educational level and fluid intelligence included as covariates. Secondly, the outcome measure (post-composite score, $Y$) was regressed on activity mode ($X$), adherence ($M$), the baseline composite score ($W$) and the interaction between adherence and baseline composite score. Again, educational level and general fluid intelligence were included as covariates.

If a significant effect of activity mode ($X$) on the mediator adherence ($M$) occurs, the moderated mediation would be expressed by a significant interaction between adherence ($M$) and baseline composite score ($W$) on the composite outcome measure ($Y$), which then would affect the indirect effect process. Subsequently, regression analyses were conducted on the mean ± 1 SD of the moderator to define the degree to which mediation varies depending on the level of the moderator. Bias-corrected (BC) bootstrapping (5000 bootstrap samples) was applied as it produces more accurate confidence intervals (CIs) (MacKinnon, Lockwood, & Williams, 2004).

For testing the hypotheses, variables were standardised to attain a common metric, and $B$ coefficients were interpreted.

### Results
Analyses were computed with complete valid data of all five tests from baseline and follow-up, not imputing missing data. In the computer condition, one woman had to be excluded due to incorrect test assessment. Thus, data on cognition and adherence were used of 80 women in the exercise condition and 80 women in the computer condition ($n = 160$), and for mixed ANCOVA, 69 women of the control group were included ($N = 229$). Data of the sample characteristics are summarised in Table 1.

There were no apparent group differences regarding demographics, educational level, years of education and fluid intelligence. Regarding baseline cognitive function, the three groups did not differ significantly. Pearson correlation coefficients of adherence, composite scores of cognition and covariates are presented for all groups in Table 2.

#### Composite score of cognition: 3 × 2 mixed ANCOVA
Distribution of each variable was normal with the exception that the educational level for the intervention groups was negatively skewed. Assumption of homogeneity for composite scores was met at baseline and after six months. Covariates were not different across all three experimental groups and the assumption of homogeneity for regression slopes was met, i.e. the
Table 2. Intercorrelations of experimental groups with complete follow-up data.

<table>
<thead>
<tr>
<th></th>
<th>Exercise group (n = 80)</th>
<th>Computer group (n = 80)</th>
<th>Control group (n = 69)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1. Pre-composite score</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>2. Post-composite score</td>
<td>0.51***</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>3. General fluid intelligence</td>
<td>0.12</td>
<td>0.24*</td>
<td>–</td>
</tr>
<tr>
<td>4. Educational level</td>
<td>0.27*</td>
<td>0.40***</td>
<td>0.18</td>
</tr>
<tr>
<td>5. Adherence</td>
<td>–0.01</td>
<td>0.13</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Note: *p < 0.05, **p < 0.01, ***p < 0.001.

Figure 2. Result of 3 × 2 mixed ANCOVA: comparing composite scores at baseline and after 6-month intervention of all three groups, controlling for educational level and general fluid intelligence.

The overall relationship between pre- and post-composite scores and the covariates were equal in all groups.

Results of the 3 × 2 mixed ANCOVA, controlling for educational level and general fluid general intelligence, showed that change of composite score from baseline to follow-up was significantly different in the intervention groups (time × group interaction, $F(2, 224) = 15.89$, $p < 0.001$). The interaction graph in Figure 2 depicts increased scores for the experimental groups and a score decline for the controls.

Results of dependent t-tests on pre- and post-composite scores (Table 3) showed a significant increase from pre- to post-testing for both experimental groups (all $p’s < 0.001$), whereas the decline of the control group was statistically significant ($p = 0.07$).

Adherence data

Both experimental groups did not differ in their average travel time (one way) to course locations: the physical exercise course reported on average 43 minutes (SD = 16) and the computer course 41 minutes (SD = 17), $t(158) = 0.85$, $p > 0.05$. However, the average course attendance was significantly higher in the computer course (85%, SD = 19) than in the physical exercise course (69%, SD = 26), with heterogeneous variances (Levene’s test $p < 0.01$). The variable adherence, defined as time spent on course participation, was homogeneous in variance for both groups, with significant differences in both groups, $F(1, 158) = 10.9$, $p = 0.001$. The computer course spent, on average, an additional 30 hours on participation as compared to the physical exercise course (exercise group: 146.5 h, SD = 63; computer group: 177 h, SD = 53).

(i) Test of mediation

Findings fit the criterion for an indirect effect but not for mediation, which is a special case of an indirect effect, as there was no significant relationship between activity mode (exercise vs. computer group) and the outcome ($B = -0.20$, $p > 0.05$). The standardised estimate of regressing adherence on the activity mode was significant ($B = 0.51$, $p = 0.001$). Adherence, controlling for activity mode, predicted cognitive change ($B = 0.21$, $p = 0.009$). Results yielded a significant indirect effect of activity mode on the cognitive composite change score through adherence (Sobel $Z = 2.01$, $p = 0.044$). Bootstrapped 95% CIs ranging from 0.02 to 0.24 corroborated the results of a mean indirect effect of $B = 0.11$, SE = 0.06.

(ii) Test of moderation

Results indicated that the baseline composite score moderated the relationship between adherence and the post-composite score ($B = -0.12$, $t = -1.84$, $p = 0.068$). This model, including the moderator and covariates, accounted for approximately 38% of the variance in the post-composite score ($R^2 = 0.38$, $p < 0.001$). We plotted the interaction at one SD above and below the mean of the post-composite score to facilitate interpretation. Women with a low-baseline composite score but high adherence are predicted to outperform women with the same baseline score but low adherence (Figure 3).

(iii) Test of moderated mediation

As the above-mentioned results suggest a moderating effect of the baseline composite score on the adherence-outcome relationship, a BC bootstrapped moderated mediation analysis was conducted to investigate the utility of the overall model. Results of the first regression analysis showed that activity mode ($X$)
predicted adherence ($M; B=0.51, p=0.001$; Table 4, mediator variable model).

The second regression analysis provided a statistical trend for the interaction between baseline composite score and adherence ($W\times M; B=−0.11, p=0.072$), suggesting that the indirect effect of activity mode ($X$) on the cognitive outcome measure ($Y$) through adherence ($M$) was moderated by the baseline composite score ($W$). It further demonstrated that the postcomposite score ($Y$) was significantly predicted by adherence ($M; B=0.17, p=0.011$) and educational level ($B=0.19, p=0.006$), but not by activity mode ($X; B=−0.20, p>0.05$) and general fluid intelligence ($B=0.05, p>0.05$; Table 4, dependent variable model). Given the statistical trend for the interaction term, significance tests were conducted on values of the mean $±1SD (0.0±1.0)$ of the moderator. Thereby, the indirect effect of activity mode on postcomposite score of cognition through adherence was seen when levels of baseline composite score on cognitive performance were low ($−1SD, p=0.023$) to moderate (mean, $p=0.057$), but not when this level was high ($+1SD, p=0.543$). BC bootstrapped CIs corroborated the results as neither of the CIs of low and mean levels of the moderator contained zero. Table 4 provides results on the bootstrapped indirect effect at the three selected levels.

The SPSS macro provided bootstrapped conditional indirect effects at different values of the moderator that fall within the range of the data using an extension of the Johnson-Neyman technique. Results showed that the conditional indirect effect was significant at $\alpha=0.05$ for any value of the centred baseline composite score smaller than $−0.13$ ($z$-standardised scores, assessed on the total sample, $N=229$, were standardised for the intervention group, $n=160$). As the baseline composite score decreased, the indirect effect became stronger. The given values of the moderator for which the indirect effect was significant delineate the region of significance, here with a lower limit of the standardised value of the baseline composite score of $−2.98$ and the higher limit of $−0.13$.

Discussion

We asked the questions of how and for whom two different intervention programmes are effective with regard to cognition. We found that spending time on a physical or mental activity is a prerequisite for cognitive improvement and that spending a high amount of time is most beneficial for participants with low-baseline cognitive performance.

We studied the relationship between adherence and cognitive performance, i.e. a composite measure of five different neuropsychological tests which allowed for distinct conclusions on the effectiveness of adherence. Composite scores are widely used in research on cognition (Helzner, Scarmeas, Cosentino, Portet, & Stern, 2007; Sturm et al., 2005; Weuve et al., 2004; Willis et al., 2006). They minimise possible ceiling effects which are likely to have played a role in our healthy and well-educated sample. The positive intervention effect on one composite measure of five cognitive tests is in line with primary analyses, where all neuropsychological tests were analysed separately (Klumsmann et al., 2010); both intervention groups improved in this composite measure over the 6 months compared to the controls who declined slightly.

It was found that the activity mode (physical vs. mental activity) predicted adherence. More specifically, the exercise group spent significantly less time for the intervention during the six months than the computer group. We conclude that physical activity is a more demanding, challenging and exhausting behaviour to perform than a mental activity with the same training characteristics. Reasons may be the most highly rated barriers to exercise reported by Newson and Kems...
Further, we asked for whom the hypothesised mediation effect occurs. The indirect effect of activity mode on cognitive performance was seen only when baseline cognitive performance was low to moderate, but not when participants performed very well (1 SD above the mean). Compared to the already high-functioning sample at baseline, low-baseline women seemed to have more room to improve or reach gains. These results are consistent with recent findings of Langbaum and collaborators (Langbaum et al., 2009), who found individuals with lower baseline memory ability to improve most in verbal learning tests.

The cognitive reserve hypothesis postulates that individuals have ‘the ability to optimize or maximize performance through differential recruitment of brain networks, which perhaps reflect the use of alternate cognitive strategies’ (Stern, 2002, p. 451). Low-baseline women in our study might have been stimulated and might have learned to use their cognitive reserve due to their enriched environment, whereas high-baseline women possibly have exploited this potential already in the beginning. For them, adherence (low or high) had the function of stabilisation and protection against cognitive decline. The decision to use an overall measure of adherence and not only the frequency with which one engages in a specific activity (Newson & Kemps, 2006) is justified by our results. We reran the performed analyses by replacing adherence with a limited measurement of adherence, i.e. the average course attendance. Interestingly, none of the results remained significant. We assume that adherence, as defined here, is an overall measure of the environmental stimulation for it includes travelling activity. For example, due to a long journey to course locations,
a woman might have had an average course attendance of only 60%. But, adding travelling time for those 60%, this would possibly compensate for lower course attendance; this woman may spend the same amount of time on participation than another woman who lives ‘around the corner’ from the location and attends almost 100% of the courses. Their cognitive improvement would be predicted to be the same. If this person had been defined as a non-adherer using a certain threshold (for details see Shields, Brawley, & Lindover, 2005), analyses on effects of an intervention would have been misleading.

We can assume that the stimulating effects continued beyond the actual intervention units: participants might have been physically stimulated while being ‘en-route’ by an additional, albeit unspecific, travel activity (Wener & Evans, 2007), because all but three participants were public transit commuters. They might as well be mentally stimulated by preparing for the ongoing courses, reworking previous course units or reflecting on various experiences and impressions which were made before, during and after the courses.

As both activity groups improved their cognitive performance compared to a control group, obviously both intervention types provided a similar ‘nutrient’ for counteracting cognitive decline. Motivation, measured by intention to participate in the courses, was high and equally strong in both courses. Translating this intention into adherence is addressed by the concept of volition (Schwarzer, 2008). Judging an individual’s effort solely in a mathematical way on course attendance is difficult, because both activity programmes are qualitatively distinct. We can only pronounce a judgement from the individual’s motivational and volitional effort and that is expressed by spending time on these activities including travel time. Future research may explore the influence of travel activity in an active control group.

Limitations
We found a linear relationship of adherence on cognitive change. A classic dose–response effect could not be calculated, because participants were not randomised to different groups of training intensity, and adherence in our study turned out to be high. In addition to the objectively measured course attendance, we used self-reported data on travel time, which may have been misjudged. Pedometers (Wener & Evans, 2007) could be used in future research to measure additional physical travelling activity. Nevertheless, this would not allow for a calculation of the temporal engagement in intervention content, such as reworking previous course units.

Conclusion
Older healthy women are able to adhere to challenging physical or mental training over an extended period of time. This enables them to improve their cognitive abilities. Results are most promising for cognitively less fit persons; they will benefit most if they invest time and effort in an intervention. The introduced adherence measure including travelling time is advantageous: it illustrates the importance of accounting for the amount of time a person is willing to spend on implementing a new challenging physical or mental activity. This temporal engagement marks a breakthrough of a person’s daily routine, which is supposed to be a prerequisite for cognitive improvement.

References


