Physically active academic lessons

de Greeff, Johannes Wilhelmus

Publication date:
2016

CHAPTER 5

Long-term effects of physically active academic lessons on physical fitness and executive functions in primary school children[5]

Johannes W de Greeff, Esther Hartman, Marijke J Mullender-Wijnsma, Roel J Bosker, Simone Doolaard & Chris Visscher.

[5]Published in: Health Education Research, 2016, 31(2):185-194.

ABSTRACT

Integrating physical activity into the curriculum has potential health and cognitive benefits in primary school children. The aim of this study was to investigate the effects of physically active academic lessons on cardiovascular fitness, muscular fitness and executive functions. In the current randomized controlled trial, 499 second and third graders within 12 primary schools (mean age = 8.1 ± 0.7) were randomized to the intervention (n = 249) or control condition (n = 250). The physically active academic lessons were given for 2 consecutive school years, 22 weeks per year, three times a week, with a duration of 20–30 min per lesson. Multiple tests were administered before, between and after the intervention period, measuring cardiovascular fitness, muscular fitness and executive functions. Multilevel analysis accounted for the nested structure of the children within classes and schools. Results showed a larger improvement in speed-coordination [B = -0.70, p = 0.002] and a lower improvement in static strength [B = -0.92, p < 0.001] for the intervention group compared with the control group. The current lessons did not result in a significant change in executive functions.

INTRODUCTION

Reducing sedentary behavior and promoting regular participation in physical activity across the school day are essential goals to prevent chronic health conditions in children, including hypertension, insulin resistance and blood lipids (Bailey et al., 2012; Boddy et al., 2014). Emerging evidence has been collected showing that physical activity also has beneficial effects on cognitive performance in children (Donnelly & Lambourne, 2011; Tomporowski et al., 2015). Despite these benefits, Dutch children spend large amounts of their school-time (66%) in sedentary activities (van Stralen et al., 2014). Changing this pattern is difficult, given the concern in many primary schools over meeting the goals for academic achievement (Donnelly et al., 2013). An innovative strategy to reduce sedentary behavior and promote physical activity during school time is integrating physically active academic lessons into the curriculum (Norris et al., 2015). These types of educational lessons aim to incorporate physical activities with a moderate-to-vigorous intensity into the teaching of academic lesson content and do not come with the cost of academic instruction time.

Although research on the effects of physically active academic lessons is still in its infancy, it is already known that children that participate in moderate-to-vigorous physical activity (MVPA) regularly, improve their cardiovascular fitness (Kriemler et al., 2010; Kristensen et al., 2010). Considering the wide variety of physical exercises provided in physically active academic lessons, including lunges and squats, these lessons might also improve muscular fitness. This can result in additional physical health benefits, such as an improved bone health (Ortega et al., 2008; Smith et al., 2014).

The benefits of integrating physical activity into the curriculum might even extend beyond the physical health benefits. Emerging evidence indicates that physical activity is beneficial for academic achievement (Donnelly & Lambourne, 2011; Mullender-Wijnsma et al., 2015a) and executive functions (EF) in children (Best, 2010; Guiney & Machado, 2013; Sibley & Etnier, 2003; Verburgh et al., 2014). EF are those cognitive processes that are needed for complex, goal-oriented cognition and behavior (Welsh et al., 2006) and consists of at least three related domains: inhibition, which is the ability to avoid dominant or automatic responses and suppressing environmental interference; working memory, which involves temporarily keeping relevant information and further processing this information; and cognitive flexibility, which require children to switch between multiple tasks (Diamond et al., 2007; Miyake et al., 2000). Previous studies have shown that it is possible to improve EF in primary school children through a 9-month physical activity program, including at least 70 min of daily MVPA (Castelli et al., 2011; Hillman et al., 2014; Kamijo et al., 2011). The rationale for physical activity to improve EF in the long term is that regular participation in physical activity of moderate-to-vigorous intensity results in improvements in physical fitness (Kriemler et al., 2010; Kristensen et al., 2010), leading to several physiological changes. Physical activity is thought to enhance the angiogenesis and neurogenesis in areas of the brain that support memory and learning (Chaddock et al., 2011), which in turn can increase the perfusion capacity in the brain, subsequently enhancing EF (Querido & Sheel, 2007; Verburgh et al., 2014). For example, an increase in physical fitness is found to be associated with greater gray and white matter volume in the hippocampus and basal ganglia (Chaddock, Erickson, Prakash, VanPatter et al., 2010; Chaddock, Erickson, Prakash, Kim et al., 2010). To date, however, no studies have reported effects of physically active academic lessons on EF.

The aim of this study was to examine the effects of physically active academic lessons on cardiovascular fitness, muscular fitness and EF after 2 years. It was hypothesized that physically active academic lessons improve physical fitness and EF. This study is part of the project 'Fit en Vaardig op school' (Fit and academically proficient at school; F&V), a cluster-randomized controlled trial (RCT), in which primary school children were randomly assigned to a control group or a 22-week classroom-based intervention program. A relatively large pilot-study in six schools has shown that the F&V lessons can be successfully implemented in the curriculum and improve academic performance in third grade children (Mullender-Wijnsma et al., 2015a). Sub-analysis of the data collected in the first intervention year of the cluster-RCT showed an average MVPA engagement of 60% (Mullender-Wijnsma et al., 2015b).

METHODS

Participants and study design

Based upon the results of the physical activity across the curriculum (PAAC) study, an effect size of 0.44 was expected (J. E. Donnelly and J. L. Greene, personal communication) (Donnelly et al., 2009). Assuming an average class size of 25 children, power analysis [power = 0.80; α = 0.05; 1-tailed; Intraclass correlation = 0.10] resulted in a required sample size of 10 intervention classes and 10 control classes (Spybrook & Raudenbush, 2008). Twelve primary schools in the Northern part of the Netherlands participated in the 2-year intervention program. A second or third grade class from each school was randomly assigned to serve as an intervention group. All children from that class participated in the intervention program. The class that was not assigned to the intervention group was automatically classified as the control group. In total, data from 499 primary school children were obtained. The mean (SD) age of the study population was 8.1 (0.7) years and included 226 boys (45.3%) and 273 girls (54.7%). After randomization, 249 children were in the intervention group and 250 in the control group (Table 5.1). The control group consisted of a higher percentage of third grade children [$\chi^2(1) = 5.22$; p = 0.025] and was significantly older [$t(497) = 2.24$; $p = 0.026$] due to a difference in number of children within each class. No significant age differences were found when analyzing the second and third grade children separately. The intervention and control group were comparable on all other descriptive characteristics. In Figure 5.1 the number of participants are shown during each measurement period. Common reasons for not completing the tests were absence from school or leaving to attend another school. Due to circumstances not related to the intervention, two schools did not start the second intervention period, resulting in a lower sample size at T2 for both the control and intervention group. A loss of two schools was taken into account during the power analysis. Informed consent was obtained for all children and all procedures were approved by the institutional Ethics Committee of the Center for Human Movement Sciences of the University Medical Center Groningen, University of Groningen.

Table 5.1 Baseline demographics of the included study population.

	Intervention n = 249	Control n = 250	*p* value
Age, years (sd)	8.0 (0.7)	8.2 (0.7)	0.026[a]
Height, cm (sd)	131.9 (7.2)	132.5 (7.1)	0.371[a]
Weight, kg (sd)	29.9 (7.4)	29.9 (6.7)	0.971[a]
Grade, n third (%)	108 (43.4)	134 (53.6)	0.025[b]
Sex, n boys (%)	116 (46.6)	110 (44.0)	0.590[b]
BMI, kg/m^2 (sd)	17.0 (2.8)	16.9 (2.6)	0.577[a]
Overweight, n (%)[c]	40 (17.2)	43 (17.8)	0.616[b]
Obesity, n (%)[c]	20 (8.6)	15 (6.2)	

[a]Independent t-test. [b]Non-parametric Chi-square test. [c]According to the reference values by Cole et al. (2012).

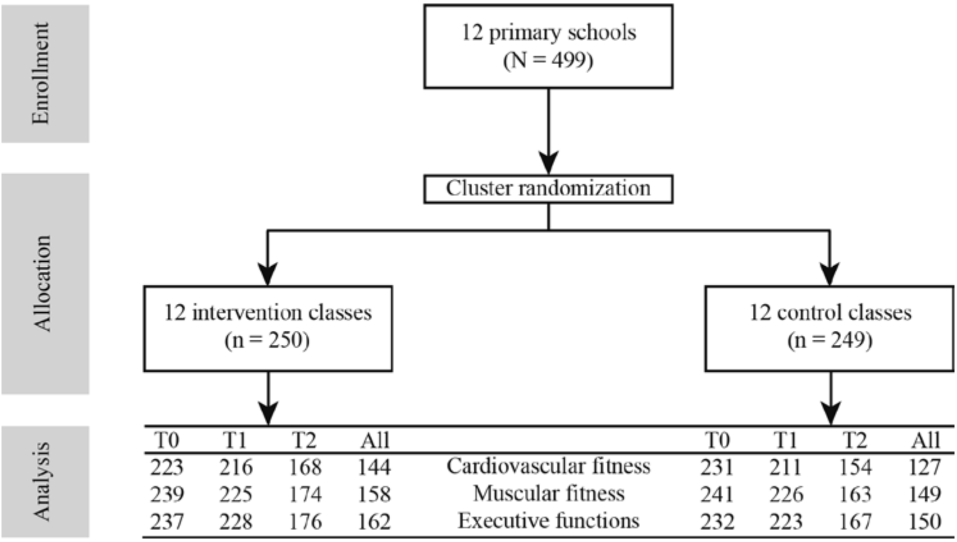

Figure 5.1 Flow chart showing the number of participants in each part of the study.

Measurements

Cardiovascular and muscular fitness

Physical fitness consists of at least two related domains: cardiovascular fitness, the capacity of the cardiovascular and respiratory system to perform continuous strenuous physical exercise; and muscular fitness, the capacity to perform work against a resistance (Ortega et al., 2008). The 10 × 5 m shuttle run (SR) (10 × 5 m SR, speed-coordination, in seconds) and the 20 m endurance SR (20 m SR, cardiorespiratory endurance, in number of completed stages) were administered for measuring cardiovascular fitness (de Greeff et al., 2014). Standing broad jump (SBJ, explosive strength, in cm), sit-ups in 30 s (SUP, muscular endurance, in number of completed SUP) and handgrip strength (HG, static strength, in kg) were administered for measuring muscular fitness (de Greeff et al., 2014). All tests were part of the standardized Eurofit test battery, which has been designed for the assessment of health-related fitness in both children as well as adults (Adam et al., 1988). Test-retest reliability [*r* between 0.62 and 0.97] and construct validity of all tests are adequate (van Mechelen et al., 1991).

Executive functions

The Golden Stroop test was used to measure inhibition (Strauss et al., 2006). In the first condition, children had to read aloud the words from a word card (color words printed in black ink). In the second (congruent) condition, children had to name the colors of colored rectangles from a color card (solid rectangles printed in red, green, yellow or blue ink). In the final (incongruent) condition, children had to name the color of the ink in which the words were written (color words printed in red, green, yellow or blue ink in which the ink does not match the word). For each card, the time was limited to 45 s and the score ranged from 0 to 100 correctly named words or colors. The Stroop interference score was calculated by subtracting the score of the incongruent condition from the score of the congruent condition (Strauss et al., 2006). The test re-test reliability [$r = 0.81$] indicate that the Stroop interference score has a good reliability within children (Neyens & Aldenkamp, 1997).

For working memory, the Digit span backward and Visual span backward were used (Wechsler, 1987). The Digit span backward involved presenting a sequence of spoken digits by the instructor followed by asking the child to recall the sequence in a reverse order. The number of digits in a sequence increased from two to eight. The test stopped when the child failed to recall at least two out of the three sequences within a span. The score ranged from 0 to 21 correctly recalled sequences. One practice trial was given to ensure that the child understood the test. The Visual span backward involved tapping a sequence of squares by the instructor on a card containing eight printed squares, followed by asking the child to tap the same squares in a reverse order. The number of tapped squares in a sequence increased from two to seven and the test stopped when the child fails to recall two sequences with the same length. The score was the total correctly tapped sequences and ranged from 0 to 12. Test re-test reliability for the Digit span backward [$r = 0.82$] and Visual span backward [$r = 0.75$] indicate that both tests are reliable (Wechsler, 1987). In addition, the factor loadings with general working memory indicate that both the Digit span backward [loading = 0.75] and the Visual memory span [loading = 0.65] are a valid measure for working memory (Wechsler, 1987).

Cognitive flexibility was measured using a modified version of the Wisconsin card sorting test (M-WCST), which is more suitable for children compared with the regular version (Cianchetti et al., 2007). The instructor placed four stimulus cards on a table in front of the child. Each stimulus card contained a unique shape (triangle, cross, circle or star), a unique color (yellow, red, green or blue) and a unique number of shapes (one, two, three or four). The child was instructed to sort the 48 response cards according to one of the unique characteristics. The child was told whether the card was sorted correctly or incorrectly. Only one sorting rule applied, after six consecutive correct responses the sorting rule changed and the instructor instructed the child to find another rule. The test stopped when all 48 cards were sorted, or when the child sorted the cards according to six different rules. A categorizing efficiency score was calculated (Cianchetti et al., 2007), meaning that for every correctly sorted rule six points were awarded and one point for each of the 48 cards not used.

Intervention program

The intervention program had a duration of 2 school years (2012–13 and 2013–14), 22 weeks per year (October until May), and consisted of three F&V lessons per week. The F&V lessons had a duration of 20–30 min, with 10–15 min spend on solving mathematical problems and 10–15 min spend on language. During the school holidays the lessons were not continued. Each F&V lesson was supported by a PowerPoint presentation and a manual describing the tasks in detail. In the first year, six primary school teachers were hired and trained to deliver the lessons. In the second year, the regular classroom teacher received a training in delivering F&V lessons and was made responsible for delivering the lessons. All F&V lessons were cognitively matched with the development of the children at each grade level. The physical activities were aimed to be of moderate-to-vigorous intensity. During the F&V lessons all children started with performing a basic exercise, such as jogging, hopping in place or marching. A specific exercise was performed when the children solved an academic task. For example, for mathematics, children had to jump eight times to solve the multiplication '4 × 2'. For language, children had to perform a squat for every spelled letter in the word 'dog'. After performing the specific exercise, children had to continue performing the basic exercise until the next academic task was shown. Heart rate monitoring results of the F&V lessons during the first intervention year of the cluster-RCT showed an average MVPA engagement of 60%, which can be translated in 14 min of MPVA per lesson (Mullender-Wijnsma et al., 2015b). A more detailed description of the intervention program and its implementation can be found elsewhere (Mullender-Wijnsma et al., 2015a; Mullender-Wijnsma et al., 2015b).

Procedure

In October 2012 (T0), EF were individually assessed in a quiet room at the school by instructed researchers. Approximately 2 weeks later, muscular and cardiovascular fitness were assessed during two regular physical education classes. Each instructed researcher received a 2 h training to get familiar with the EF and physical fitness tests and were mostly blinded to the condition children had been allocated to (during 88.6% of the measurements). The 12 intervention classes followed the intervention program directly after the baseline measurements, while the 12 control classes followed the regular lessons. In May 2013, 6 months after the start of the first intervention period (T1) and in May 2014 (directly after the second intervention period; T2), the children were tested using the same EF and physical fitness tests as was used during T0. During T1 and T2, assessments started with muscular and cardiovascular fitness to ensure that the physical fitness tests were administered as close to the start and end of the intervention program as possible.

Statistical analysis

Preliminary descriptive statistics were performed using SPSS for Windows (Version 22.0). Independent *t*-tests were performed on each test of cardiovascular fitness, muscular fitness and EF, comparing the scores of intervention group with the control group at T0, T1 and T2.

Multilevel regression analyses (MLwiN, version 2.29) were conducted for the main analysis. Multilevel analysis is particularly appropriate when using data with a nested source of variability, which is the case when studying children within multiple classes and schools (Snijders & Bosker, 2011). The outcome measures of the cardiovascular fitness tests (10 × 5 m and 20 m SR), muscular fitness tests (SBJ, SUP and HG) and the EF tests (Stroop test, Digit span, Visual span and M-WCST) at T2 were used as the dependent variables. For each dependent variable, three models were built in order to investigate the effects of the intervention on cardiovascular fitness, muscular fitness and EF. The covariates model contained only measurement period (categorical; T0, T1 and T2), grade and sex as fixed effects. Model 1 contained condition (intervention or control) as an additional fixed effect. Model 1 was used to investigate whether the intervention group differed from the control group after 2 years. Model 2 contained condition and the interaction between condition and measurement period (Condition × T1 and Condition × T2) as additional fixed effects. Model 2 was used to investigate whether the possible effects after 1 year were different from the effects after 2 years. For each model, school (level 3), child (level 2) and measurement period (level 1) were treated as random effects. The model fit was evaluated by comparing the deviance of the covariates model with the deviance of Model 1 and 2. Statistical significance was adopted for all tests when $p < 0.05$.

RESULTS

Descriptive statistics

Table 5.2 shows the results of the cardiovascular fitness, muscular fitness and EF outcomes at T0, T1 and T2. Children in the control group scored significantly higher on HG at T2 [$t(341) = 2.87$; $p = 0.004$] compared with the intervention group. For the 10 × 5 m SR, the intervention group performed significantly lower compared with the control group at T0 [$t(480) = -2.31$; $p = 0.021$) and at T1 [$t(453) = -2.55$; $p = 0.011$]. No significant differences between the control and intervention group were found for SBJ, SUP and the 20 m SR. Also, no significant differences were found for the tests of EF.

Effects on cardiovascular and muscular fitness

Table 5.3 shows the effects of the intervention (Model 2) on the tests of cardiovascular and muscular fitness. For HG, Model 2 significantly improved the model fit compared with the covariates model [$\Delta\chi^2(3) = 17.1$; $p < 0.001$]. T2 [$B = 3.26$; $p < 0.001$] and the interaction Condition × T2 were significant [$B = -0.92$; $p < 0.001$], indicating that HG has improved at T2, but the improvement for the intervention group [$B = 2.36$] was significantly smaller than the control group [$B = 3.26$]. No differences between the groups were found at T0 and T1. For the 10 × 5 m SR, Model 2 showed a significant improvement in model fit compared with the covariates model [$\Delta\chi^2(3) = 17.5$; $p < 0.001$]. A significant effect was found for T2 [$B = -0.85$; $p < 0.001$], Condition [$B = 0.65$; $p = 0.002$] and the interaction Condition × T2 [$B = -0.70$; $p = 0.002$]. This indicates that at T0, the intervention group [$B = 0.65$] was significantly slower than the control group. At T2, scores for the intervention group [$B = -0.90$] did not differ significantly from the control group [$B = -0.85$]. In other words, the intervention group demonstrated a larger improvement on the 10 × 5 m SR after 2 years, compared with the control group. As their performance at T0 was significantly lower, the intervention group did not score significantly higher at T2. For 20 m SR, SBJ and SUP, adding condition to the model (Model 1), or adding the interaction between condition and the intervention period (Model 2) did not significantly improve the covariates model.

Effects on EF

For EF, neither adding condition to the model (Model 1), nor adding the interactions between condition and intervention period (Model 2) did significantly improve the covariates model for any of the tests (Table 5.4). This indicates that there were no differences between the intervention and control group during T0, T1 and T2.

Table 5.2 Mean scores on the domains of cardiovascular fitness, muscular fitness and executive functions at T0, T1 and T2, separated for the intervention and control group.

		Intervention n = 249	Control n = 250
Cardiovascular fitness			
10x5m shuttle run, s	T0	24.8 (2.6)*	24.3 (2.7)*
	T1	25.0 (2.4)*	24.4 (2.6)*
	T2	23.3 (2.3)	23.4 (2.6)
20m shuttle run, stage	T0	3.9 (1.7)	4.1 (2.0)
	T1	4.3 (1.8)	4.2 (2.0)
	T2	4.7 (2.0)	4.9 (2.2)
Muscular fitness			
Standing broad jump, cm	T0	123.0 (20.1)	125.5 (21.3)
	T1	124.1 (18.8)	123.5 (19.5)
	T2	129.7 (21.4)	131.2 (19.9)
Sit-ups, n	T0	15.1 (3.9)	15.4 (4.2)
	T1	16.4 (4.7)	16.3 (4.8)
	T2	16.7 (4.9)	17.2 (4.7)
Handgrip strength, kg	T0	13.3 (3.3)	13.5 (3.3)
	T1	14.9 (3.7)	15.0 (3.5)
	T2	15.8 (3.7)**	17.0 (3.8)**
Executive functions			
Stroop	T0	17.9 (7.9)	17.5 (8.1)
	T1	18.4 (8.1)	19.0 (7.5)
	T2	19.6 (8.1)	19.9 (9.5)
Digit span backward	T0	5.1 (1.7)	5.0 (1.5)
	T1	5.6 (1.7)	5.6 (1.7)
	T2	6.0 (2.2)	6.2 (1.9)
Visual span backward	T0	5.6 (1.9)	5.7 (1.8)
	T1	6.5 (1.7)	6.7 (1.7)
	T2	6.6 (1.7)	6.8 (1.6)
M-WCST	T0	20.8 (11.4)	21.7 (11.8)
	T1	26.8 (12.1)	28.0 (12.5)
	T2	31.6 (10.8)	30.3 (11.3)

Values are mean (SD).
M-WCST = modified Wisconsin card sorting test.
*Significant differences between control and intervention group with $p < 0.05$;**Significantly different at $p < 0.01$.

Table 5.3 Results of the multilevel analyses for cardiovascular and muscular fitness (Model 2).

	Cardiovascular fitness		Muscular fitness		
	10x5m SR	20m SR	SBJ	SUP	HG
Fixed effects (SE)					
Intercept	25.00 (0.31), $p < 0.001$	3.44 (0.21), $p < 0.001$	117.21 (1.62), $p < 0.001$	14.17 (0.41), $p < 0.001$	11.90 (0.28), $p < 0.001$
Grade[a]	-0.96 (0.18), $p < 0.001$	0.24 (0.14), $p = 0.095$	9.18 (1.52), $p < 0.001$	1.45 (0.34), $p < 0.001$	2.49 (0.27), $p < 0.001$
Sex[b]	-0.89 (0.18), $p < 0.001$	1.14 (0.14), $p < 0.001$	7.95 (1.52), $p < 0.001$	1.00 (0.34), $p = 0.003$	0.69 (0.27), $p = 0.012$
T1	0.16 (0.15), $p = 0.265$	0.19 (0.11), $p = 0.078$	-2.36 (1.12), $p = 0.035$	0.92 (0.28), $p < 0.001$	1.56 (0.16), $p < 0.001$
T2	-0.85 (0.17), $p < 0.001$	0.87 (0.12), $p < 0.001$	4.56 (1.26), $p < 0.001$	1.73 (0.31), $p < 0.001$	3.26 (0.17), $p < .001$
Condition[c]	**0.65 (0.21), $p = 0.002$**	-0.08 (0.16), $p = 0.613$	-1.86 (1.74), $p = 0.283$	-0.16 (0.40), $p = 0.685$	0.02 (0.30), $p = 0.954$
Condition*T1	-0.05 (0.21), $p = 0.804$	0.12 (0.15), $p = 0.305$	3.29 (1.58), $p = 0.038$	0.33 (0.39), $p = 0.402$	-0.05 (0.22), $p = 0.806$
Condition*T2	**-0.70 (0.23), $p = 0.002$**	-0.25 (0.17), $p = 0.138$	2.16 (1.76), $p = 0.220$	-0.33 (0.43), $p = 0.450$	**-0.92 (0.24), $p < 0.001$**
Model summary					
Deviance statistics	5446.6	4594.4	10710.5	7034.8	5949.7
Number of estimated parameters	11	11	11	11	11

Significant effects related to the hypotheses are shown in bold font.
10 × 5 m SR = 10 × 5 m shuttle run, 20 m SR = 20 m endurance shuttle run, SBJ = standing broad jump, SUP = sit-ups, HG = handgrip strength.
[a,b,c]Respectively second grade, girls and control group are coded as 0. Third grade, boys and intervention group are coded as 1.

Table 5.4 Results of the multilevel analyses for executive functions (Model 2).

	Stroop	Digit span	Visual span	M-WCST
Fixed effects (SE)				
Intercept	16.40 (0.64), $p < 0.001$	4.60 (0.14), $p < 0.001$	5.21 (0.15), $p < 0.001$	19.54 (0.99), $p < 0.001$
Grade[a]	1.57 (0.54), $p = 0.004$	0.76 (0.12), $p < 0.001$	0.63 (0.12), $p < 0.001$	5.13 (0.85), $p < 0.001$
Sex[b]	0.61 (0.54), $p = 0.259$	-0.04 (0.12), $p = 0.751$	0.34 (0.12), $p = 0.005$	-1.39 (0.85), $p = 0.102$
T1	1.52 (0.65), $p = 0.020$	0.58 (0.13), $p < 0.001$	0.96 (0.13), $p < 0.001$	6.26 (0.78), $p < 0.001$
T2	2.28 (0.72), $p = 0.002$	1.12 (0.14), $p < 0.001$	1.07 (0.14), $p < 0.001$	9.11 (0.87), $p < 0.001$
Condition[c]	0.57 (0.74), $p = 0.443$	0.14 (0.16), $p = 0.383$	-0.05 (0.16), $p = 0.767$	-0.42 (1.04), $p = 0.685$
Condition*T1	-1.06 (0.92), $p = 0.250$	-0.06 (0.18), $p = 0.732$	-0.08 (0.18), $p = 0.672$	-0.09 (1.09), $p = 0.937$
Condition*T2	-0.83 (1.01), $p = 0.408$	-0.27 (0.20), $p = 0.171$	-0.14 (0.20), $p = 0.469$	1.55 (1.21), $p = 0.200$
Model summary				
Deviance statistics	9035.5	4981.9	4966.2	9760.3
Number of estimated parameters	11	11	11	11

M-WCST = Modified Wisconsin card sorting test.
[a,b,c]Respectively second grade, girls and control group are coded as 0. Third grade, boys and intervention group are coded as 1.

DISCUSSION

The aim of this study was to examine the effects of physically active academic lessons on cardiovascular fitness, muscular fitness and EF after 2 years. The results showed that, for cardiovascular fitness, a larger improvement in speed-coordination (assessed with the 10 × 5 m SR) was found in the intervention group, compared with the control group. For muscular fitness, a smaller improvement in static strength (assessed with HG) was found in the intervention group compared with the control group. For EF, no significant differences were found between the control and intervention group after 1 or 2 years.

Although the intervention group had significantly lower baseline scores, the increase in speed-coordination after 2 intervention years was larger for the intervention group compared with the control group. This larger increase for the intervention group could be due to the short intermitted nature of most of the physical activities during the physically active academic lessons. For example, the children had to spell a word by jumping in place for every mentioned letter. After spelling the word, children had to switch to jogging in place until the next word was shown. The intervention program did not result in an improvement of cardiorespiratory endurance (assessed with the 20 m SR) after 2 intervention years. A previous study has shown that a 9-month physical activity program, which provided first and fifth graders about 13 min of additional in-school MVPA on a daily basis, resulted in a significant improvement of cardiovascular fitness (Kriemler et al., 2010). The children in this study engaged in similar amounts of MVPA per lesson, but the lessons were provided three times a week instead of daily. It is possible that the frequency of the physically active academic lessons was not enough to positively affect the scores of both tests measuring cardiovascular fitness. Several other studies recommended at least 30 min of daily MVPA to improve cardiovascular fitness in school-aged children (Hillman et al., 2014; Janssen & LeBlanc, 2010; Kamijo et al., 2011). We therefore hypothesize that future studies may need to focus on interventions programs that provide MPVA on a daily basis in order to positively affect cardiovascular fitness.

It is widely assumed that physical activity has a positive effect on EF in children (Best, 2010; Guiney & Machado, 2013; Sibley & Etnier, 2003; Tomporowski et al., 2015; Verburgh et al., 2014). A recent meta-analysis, investigating the effects of physical activity on EF, showed a moderate positive effect size for studies that measured EF directly (within minutes or hours) after a single short-term exercise bout [$d = 0.57$] (Verburgh et al., 2014). However, as a result of insufficient studies with a high-quality design, it is difficult to make conclusions about the effect of long-term physical activity interventions on EF (Chaddock et al., 2011; Guiney & Machado, 2013). This study showed that the children did not significantly improve on the four tests representing EF after 2 years. In contrast, a 9-month after-school physical activity program (FIT Kids program), including at least 70 min of MVPA per day, resulted in improvements in working memory (Hillman et al., 2014; Kamijo et al., 2011), inhibition and cognitive flexibility (Hillman et al., 2014). In addition, children participating in the FIT kids program improved their cardiovascular fitness (Hillman et al., 2014). A possible explanation for these conflicting results is that the effect of physical activity on EF could partly be mediated by cardiovascular fitness.

The cardiovascular fitness hypothesis states that changes in cardiovascular fitness, as a result of regular MVPA, leads to physiological changes in the brain structural network (Etnier et al., 2006). Examples of these physiological changes are an elevated cerebral blood flow and an increased concentration level of growth factors responsible for neurogenesis and the synaptic plasticity of the brain (Dishman et al., 2006; Hillman et al., 2008). The current intervention program, in contrast with the FIT Kids program, was unable to improve cardiovascular fitness, which might partly explain the lack of significant effects in the present study. However, more research on the mediating role of cardiovascular fitness in this age group needs to be undertaken before the effects of physical activity on EF is clearly understood.

Some limitations should be considered for future research. First, despite randomly assigning the classes to either the control or intervention group, age and the number of second and third graders differed at baseline. As no age differences were found when analyzing the second and third grade children separately, the significant differences in grade likely resulted in the significant age difference. To control for these differences, grade was added as a control variable in the multilevel models. In addition, the intervention group performed lower on speed-coordination at baseline, compared with the control group. Although we controlled for differences at baseline in the multilevel models, the results need to be interpreted with caution. Second, it remains unclear whether or not physically active academic lessons affected the physical activity during out-of-school time. According to the 'activitystat hypothesis', children compensate the increased physical activity during one part of the day by a decrease in physical activity during another part of the day in order to maintain a particular physical activity set point (Ridgers et al., 2014; Rowland, 1998). It is therefore possible that the children in the intervention group were less active during out-of-school time because of the increase in physical activity during in-school time. Future intervention studies should therefore measure the time spend in out-of-school and in-school MVPA. Last, in the second intervention year the regular classroom teacher delivered the intervention instead of the hired teacher. Although the effects of the first intervention year did not differ from the second intervention year, this change in teacher could have confounded the efficacy of the delivery of the intervention. Strengths of this study were the RCT design and the large sample size.

In conclusion, our program provided a unique approach as it integrated physical activity into the teaching of academic lesson content. This study is, according to the authors knowledge, the first to report long-term effects of physically active academic lessons on cardiovascular fitness, muscular fitness and/or EF. Effects on cardiovascular and muscular fitness were small, but the intervention program resulted in an improvement in speed-coordination. The current lessons did not result in a significant change in EF. These results again highlight the difficulty to positively influence some of the health and cognitive aspects in primary school children. Our findings nevertheless provide important practical considerations for future studies that focus on physically active academic lessons.

REFERENCES

Adam, C., et al. (1988). *EUROFIT: European test of physical fitness*. Rome: Council of Europe, Committee for the Development of Sport.

Bailey, D. P., Boddy, L. M., Savory, L. A., Denton, S. J., & Kerr, C. J. (2012). Associations between cardiorespiratory fitness, physical activity and clustered cardiometabolic risk in children and adolescents: The HAPPY study. *European Journal of Pediatrics, 171*(9), 1317-1323.

Best, J. R. (2010). Effects of physical activity on children's executive function: Contributions of experimental research on aerobic exercise. *Developmental Review, 30*(4), 331-351.

Boddy, L. M., Murphy, M. H., Cunningham, C., Breslin, G., Foweather, L., Gobbi, R., et al. (2014). Physical activity, cardiorespiratory fitness, and clustered cardiometabolic risk in 10- to 12-year-old school children: The REACH Y6 study. *American Journal of Human Biology : The Official Journal of the Human Biology Council, 26*(4), 446-451.

Castelli, D. M., Hillman, C. H., Hirsch, J., Hirsch, A., & Drollette, E. (2011). FIT kids: Time in target heart zone and cognitive performance. *Preventive Medicine, 52*, S55-S59.

Chaddock, L., Erickson, K. I., Prakash, R. S., Kim, J. S., Voss, M. W., VanPatter, M., et al. (2010). A neuroimaging investigation of the association between aerobic fitness, hippocampal volume, and memory performance in preadolescent children. *Brain Research, 1358*, 172-183.

Chaddock, L., Erickson, K. I., Prakash, R. S., VanPatter, M., Voss, M. W., Pontifex, M. B., et al. (2010). Basal ganglia volume is associated with aerobic fitness in preadolescent children. *Developmental Neuroscience, 32*(3), 249-256.

Chaddock, L., Pontifex, M. B., Hillman, C. H., & Kramer, A. F. (2011). A review of the relation of aerobic fitness and physical activity to brain structure and function in children. *Journal of the International Neuropsychological Society, 17*(06), 975-985.

Cianchetti, C., Corona, S., Foscoliano, M., Contu, D., & Sannio-Fancello, G. (2007). Modified Wisconsin card sorting test (MCST, MWCST): Normative data in children 4–13 years old, according to classical and new types of scoring. *The Clinical Neuropsychologist, 21*(3), 456-478.

Cole, T., & Lobstein, T. (2012). Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. *Pediatric Obesity, 7*(4), 284-294.

De Greeff, J. W., Hartman, E., Mullender-Wijnsma, M. J., Bosker, R. J., Doolaard, S., & Visscher, C. (2014). Physical fitness and academic performance in primary school children with and without a social disadvantage. *Health Education Research, 29*(5), 853-860.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318*(5855), 1387-1388.

Dishman, R. K., Berthoud, H., Booth, F. W., Cotman, C. W., Edgerton, V. R., Fleshner, M. R., et al. (2006). Neurobiology of exercise. *Obesity, 14*(3), 345-356.

Donnelly, J. E., Greene, J., Gibson, C. A., Sullivan, D. K., Hansen, D. M., Hillman, C. H., et al. (2013). Physical activity and academic achievement across the curriculum (A + PAAC): Rationale and design of a 3-year, cluster-randomized trial. *BMC Public Health, 13*, 307.

Donnelly, J. E., & Lambourne, K. (2011). Classroom-based physical activity, cognition, and academic achievement. *Preventive Medicine, 52*, S36-S42.

Donnelly, J. E., Greene, J. L., Gibson, C. A., Smith, B. K., Washburn, R. A., Sullivan, D. K., et al. (2009). Physical activity across the curriculum (PAAC): A randomized controlled trial to promote physical activity and diminish overweight and obesity in elementary school children. *Preventive Medicine, 49*(4), 336-341.

Etnier, J. L., Nowell, P. M., Landers, D. M., & Sibley, B. A. (2006). A meta-regression to examine the relationship between aerobic fitness and cognitive performance. *Brain Research Reviews, 52*(1), 119-130.

Guiney, H., & Machado, L. (2013). Benefits of regular aerobic exercise for executive functioning in healthy populations. *Psychonomic Bulletin & Review, 20*(1), 73-86.

Hillman, C. H., Erickson, K. I., & Kramer, A. F. (2008). Be smart, exercise your heart: Exercise effects on brain and cognition. *Nature Reviews Neuroscience, 9*(1), 58-65.

Hillman, C. H., Pontifex, M. B., Castelli, D. M., Khan, N., Raine, L., Scudder, M., et al. (2014). Effects of the FITKids randomized controlled trial on executive control and brain function. *Pediatrics, 134*(4), e1063-e1071.

Janssen, I., & LeBlanc, A. G. (2010). Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *International Journal of Behavioral Nutrition and Physical Activity, 7*(40), 1-16.

Kamijo, K., Pontifex, M. B., O'Leary, K. C., Scudder, M. R., Wu, C., Castelli, D. M., et al. (2011). The effects of an afterschool physical activity program on working memory in preadolescent children. *Developmental Science, 14*(5), 1046-1058.

Kriemler, S., Zahner, L., Schindler, C., Meyer, U., Hartmann, T., Hebestreit, H., et al. (2010). Effect of school based physical activity programme (KISS) on fitness and adiposity in primary schoolchildren: Cluster randomised controlled trial. *BMJ, 340*, c785.

Kristensen, P. L., Moeller, N. C., Korsholm, L., Kolle, E., Wedderkopp, N., Froberg, K., et al. (2010). The association between aerobic fitness and physical activity in children and adolescents: The European Youth Heart Study. *European Journal of Applied Physiology, 110*(2), 267-275.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100.

Mullender-Wijnsma, M. J., Hartman, E., de Greeff, J. W., Bosker, R. J., Doolaard, S., & Visscher, C. (2015a). Improving academic performance of school-age children by physical activity in the classroom: Year one program evaluation. *Journal of School Health, 85*(6), 365-371.

Mullender-Wijnsma, M. J., Hartman, E., de Greeff, J. W., Bosker, R. J., Doolaard, S., & Visscher, C. (2015b). Moderate-to-vigorous physically active academic lessons and academic engagement in children with and without a social disadvantage: A within subject experimental design. *BMC Public Health, 15*(1), 404.

Neyens, L. G., & Aldenkamp, A. P. (1997). Stability of cognitive measures in children of average ability. *Child Neuropsychology, 3*(3), 161-170.

Norris, E., Shelton, N., Dunsmuir, S., Duke-Williams, O., & Stamatakis, E. (2015). Physically active lessons as physical activity and educational interventions: A systematic review of methods and results. *Preventive Medicine, 72*, 116-125.

Ortega, F. B., Ruiz, J. R., Castillo, M. J., & Sjöström, M. (2008). Physical fitness in childhood and adolescence: A powerful marker of health. *International Journal of Obesity, 32*(1), 1-11.

Querido, J. S., & Sheel, A. W. (2007). Regulation of cerebral blood flow during exercise. *Sports Medicine, 37*(9), 765-782.

Ridgers, N. D., Timperio, A., Cerin, E., & Salmon, J. (2014). Compensation of physical activity and sedentary time in primary school children. *Medicine and Science in Sports and Exercise, 46*(8), 1564-1569.

Rowland, T. W. (1998). The biological basis of physical activity. *Medicine and Science in Sports and Exercise, 30*(3), 392-399.

Sibley, B. A., & Etnier, J. L. (2003). The relationship between physical activity and cognition in children: A meta-analysis. *Pediatric Exercise Science, 15*(3), 243-256.

Smith, J. J., Eather, N., Morgan, P. J., Plotnikoff, R. C., Faigenbaum, A. D., & Lubans, D. R. (2014). The health benefits of muscular fitness for children and adolescents: A systematic review and meta-analysis. *Sports Medicine, 44*(9), 1209-1223.

Snijders, T., & Bosker, R. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage Publications Limited.

Spybrook, J., & Raudenbush, S. W. (2008). *Optimal design software [computer software].* Michigan: University of Michigan.

Strauss, E. H., et al. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. New York: Oxford University Press.

Tomporowski, P. D., McCullick, B., Pendleton, D. M., & Pesce, C. (2015). Exercise and children's cognition: The role of exercise characteristics and a place for metacognition. *Journal of Sport and Health Science, 4*(1), 47-55.

Van Mechelen, W., et al. (1991). *Eurofit: Handleiding met referentieschalen voor 12- tot en met 16-jarige jongens en meisjes in Nederland (manual with referencevalues for 12-16 year old boys and girls in the Netherlands)*. Haarlem: De Vrieseborch.

van Stralen, M. M., Yıldırım, M., Wulp, A., te Velde, S. J., Verloigne, M., Doessegger, A., et al. (2014). Measured sedentary time and physical activity during the school day of European 10- to 12-year-old children: The ENERGY project. *Journal of Science and Medicine in Sport, 17*(2), 201-206.

Verburgh, L., Königs, M., Scherder, E. J., & Oosterlaan, J. (2014). Physical exercise and executive functions in preadolescent children, adolescents and young adults: A meta-analysis. *British Journal of Sports Medicine, 48*(12), 973-979.

Wechsler, D. (1987). *WMS-R: Wechsler memory scale-revised: Manual.* San Antonio: Psychological Corporation.

Welsh, M. C., et al. (2006). Executive functions in developing children: Current conceptualizations and questions for the future. In K. McCartney, & D. Philips (Eds.), *The Blackwell handbook of early childhood development.* (pp. 167-187). Oxford: Blackwell Publishing Ltd.