

University of Groningen

Follow-Up Study Investigating the Effects of a Physically Active Academic Intervention

Mullender-Wijnsma, Marijke Johanna; Hartman, Esther; de Greeff, Johannes Wilhelmus ;
Doolaard, Simone; Bosker, Roel Jan; Visscher, Chris

Published in:
Early Childhood Education Journal

DOI:
[10.1007/s10643-019-00968-y](https://doi.org/10.1007/s10643-019-00968-y)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Mullender-Wijnsma, M. J., Hartman, E., de Greeff, J. W., Doolaard, S., Bosker, R. J., & Visscher, C. (2019). Follow-Up Study Investigating the Effects of a Physically Active Academic Intervention. *Early Childhood Education Journal*, 47(6), 699–707. <https://doi.org/10.1007/s10643-019-00968-y>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Follow-Up Study Investigating the Effects of a Physically Active Academic Intervention

Marijke Johanna Mullender-Wijnsma¹ · Esther Hartman² · Johannes Wilhelmus de Greeff² · Simone Doolaard¹ · Roel Jan Bosker¹ · Chris Visscher²

Published online: 19 August 2019
© The Author(s) 2019

Abstract

In previous studies, the effects of physically active academic lessons on academic achievement have been shown. Less is known about follow-up effects and the effects for disadvantaged groups. The first aim of this study was to examine 7–9 months follow-up effects of a physically active academic intervention on academic achievement. The second aim was to examine the effects of the 2-year intervention for a subgroup of socially disadvantaged children. A cluster-randomized controlled trial with 499 children (113 socially disadvantaged children) from second- and third-grade classes was conducted. Children's academic achievement was measured before the intervention started, after the first and second intervention year, and 7–9 months after the intervention ended. At the 7–9 months follow-up, the intervention group showed significantly greater gains in math performance in comparison with the control group. No significant follow-up effects were found on language performance. Furthermore, the lessons significantly improved the math and spelling performance of socially disadvantaged children after two intervention years. These children did not benefit more from the lessons than other children. In conclusion, effects of physically active academic lessons on math achievement persist when the lessons are no longer taught, and the lessons are an innovative way to improve the academic achievement of socially disadvantaged children. The findings suggest that physically active academic lessons should be considered for inclusion in school curriculums in order to improve the academic achievement of all children.

Trial Registration: This study is registered at www.isrctn.com (No. ISRCTN17021806).

Keywords Physically active academic lessons · Socially disadvantaged children · Academic achievement

✉ Marijke Johanna Mullender-Wijnsma
m.j.wijnsma@rug.nl

Esther Hartman
e.hartman@umcg.nl

Johannes Wilhelmus de Greeff
j.w.de.greeff@umcg.nl

Simone Doolaard
s.doolaard@rug.nl

Roel Jan Bosker
r.j.bosker@rug.nl

Chris Visscher
c.visscher@umcg.nl

¹ University of Groningen, Faculty of Behavioral and Social Sciences, Groningen Institute for Educational Research, Grote Rozenstraat 3, 9712 TG Groningen, The Netherlands

² University of Groningen, University Medical Center Groningen, Center for Human Movement Sciences, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands

Background

There is growing evidence for the positive effects of physically active academic lessons, in which physical exercises are integrated into academic lessons, on academic outcomes (Donnelly et al. 2009; Norris et al. 2015; Watson et al. 2017; Webster et al. 2015). Review studies have shown that the main goal of physically active academic interventions was to increase physical activity levels and to reduce sedentary time, but occasionally effects on academic outcomes were also reported (Norris et al. 2015; Watson et al. 2017; Webster et al. 2015). In the short term, it was found that physically active academic lessons positively influenced children's on-task behavior and academic motivation (Grieco et al. 2009; Mahar et al. 2006; Mullender-Wijnsma et al. 2015a; Vazou et al. 2012). Since academic engagement and academic motivation are important for children's academic success (Greenwood et al. 2002; Linnenbrink and Pintrich

2002) physically active academic lessons could lead to increased academic achievement in the longer term.

So far, three cluster-randomized control studies have examined the effects of prolonged physically active academic interventions. Positive effects of the Physical Activity Across the Curriculum project were found on math ($ES = 0.44$), reading ($ES = 0.35$), and spelling ($ES = 0.45$) after three intervention years (J.E. Donnelly and J.L. Greene, personal communication). In this project, a variety of academic areas were coupled with moderate to vigorous physical activity (Donnelly et al. 2009; Donnelly and Lambourne 2011). Based on these results, they recently conducted another cluster-randomized controlled trial to further examine the effects of physically active lessons on academic achievement. In this 3-year study, no effects on academic outcomes were found (Donnelly et al. 2017). We recently reported the outcomes of 'Fit & Vaardig op school' (F&V; Fit and academically proficient at school), a 2-year cluster-randomized controlled trial in the Netherlands (Mullender-Wijnsma et al. 2016). The intervention used physical activity in the teaching of math and language. After two intervention years, the children that followed the F&V lessons had significantly greater gains in math speed scores ($ES = 0.51$), general math scores ($ES = 0.42$), and spelling scores ($ES = 0.45$) in comparison with the control group. No effects were found on reading.

These findings provide evidence for the positive effects of physically active academic lessons on academic achievement. However, the need to follow the effects of physically active academic lessons over time is necessary to determine whether effects are lasting when the lessons are no longer taught. Another question that arises is whether integrating physical exercise into academic lessons also improves the academic achievement of children from disadvantaged groups. The academic achievement gap between members of low and high socio-economic status groups is a worldwide problem (Reardon 2011; Rothstein 2009). This gap also exists in the Netherlands: socially disadvantaged children (SDC) achieve worse than children without a disadvantage (non-SDC) (De Greeff et al. 2014; Driessen and Dekkers 2008). Recent studies show that despite many efforts for over 40 years, the achievement gap remains large (Driessen and Dekkers 2008). It seems that new teaching methods might be necessary to enhance the academic achievement of SDC.

The first aim of the current study was to examine what happened with the effects of the F&V intervention on academic achievement more than half a year after the end of the intervention (follow-up), when the lessons were no longer taught. Because the children could no longer benefit from the F&V lessons, we expected that their math and spelling improvements would be lower than they were after two intervention years (Mullender-Wijnsma et al. 2016). The second aim was to examine the effects

of the intervention especially for SDC. We expected that the intervention would improve the academic achievement of SDC, and because these children are more often overweight and less physically active than other children (Fredriks et al. 2005; de Vries et al. 2005), we expected that they would benefit more from the intervention than non-SDC.

Methods

Participants

The participants were 499 children from second- and third-grade classes of twelve elementary mainstream schools in the Netherlands that agreed to participate in the study. The schools were selected from 5 elementary school boards (46 schools). One hundred thirteen children were classified as SDC. The classification into SDC and non-SDC was based on parental education: the children whose parents or guardians had completed less than 3 years of secondary school were classified as SDC (Ministry of Education, Culture and Science 2006).

At each school, a second- and a third-grade class participated in the study. Per school the classes were randomly assigned to the intervention ($n = 249$) or the control group ($n = 250$). Randomization was performed by the Netherlands Bureau for Economic Policy Analysis (CPB). When the second grade was assigned to the intervention group, the third grade automatically served as control, and vice versa. Based on the initial findings from the Physical Activity Across the Curriculum study (Donnelly et al. 2009; Donnelly and Lambourne 2011), an effect size of 0.44 was assumed (J.E. Donnelly and J.L. Greene, personal communication). The power analysis resulted in a total sample of ≥ 20 classes, with 25 children per class (power 0.8, 1-tailed, $\alpha = 0.05$) (Spybrook and Raudenbush 2008).

Intervention

In the F&V intervention, physical exercise was used when teaching math and language in the classroom. In a 20–30 min lesson, half of the time was spent on math, and half on language activities. The main focus was on constant practice and repetition of concepts learned in earlier classes. The physical exercises were of moderate to vigorous intensity. For example, children spelled a word by making a squat for each letter mentioned. The lessons were supported by presentations on interactive whiteboards. A previous study

showed that the intervention program could be successfully implemented (Mullender-Wijnsma et al. 2015b).

Instruments

Academic achievement was measured using the One-Minute test (Brus and Voeten 1973) to assess reading, and the Speed-Test arithmetic (de Vos 1992) to assess the math speed performance. In addition, ability scores on math and spelling were retrieved from a child academic monitoring system, a standardized and norm-referenced test battery (Janssen and Hickendorff 2008; Janssen et al. 2010; de Wijs et al. 2010).

The One-Minute test is used to assess children's technical reading skills from second to sixth grade. The children get one minute to read aloud as many words as possible; this is then repeated with a different set of words. The score is calculated as the total number of words read correctly (from 0 to 232). For the Dutch population (based on a representative sample of 25 Dutch elementary schools), the construct validity (r varied from .78 to .86) and test-retest reliability (r varied from .89 to .92) of the test are good (Brus and Voeten 1973).

The Speed-Test arithmetic is used to assess children's math speed performance. Arithmetic problems have to be solved as quickly as possible. The total number of tasks solved determines the score (from 0 to 200). The test has been standardized in the Netherlands on a representative sample of 4804 elementary school children from 54 schools (de Vos 1992).

In the Netherlands, the spelling and math test from the child academic monitoring system is administered twice a year by most elementary schools. During the first part of the spelling test a sentence is read out by the teacher. A certain word from the sentence is then repeated, and the children have to write that word down correctly. The second part is an individual task in which the children identify misspelled words. For the Dutch population (based on a representative sample of 59 Dutch elementary schools), the reliability (r varied from .90 to .93), and content and construct validity of the spelling test were good (de Wijs et al. 2010). The math test is an individual task and involves several domains: number sense, arithmetic, algebra, geometry, time and money, and knowledge of ratios and fractions (Janssen and Hickendorff 2008); The test can be done digitally or with pencil and paper. For the Dutch population (based on a representative sample of 189 Dutch elementary schools), the reliability (r varied from .93 to .96) and content and construct validity of the test were good (Janssen et al. 2010).

Procedure

The children were assessed before the intervention (T0), after the first intervention year [8 months after T0 (T1)], after the second intervention year [1 year after T1 (T2)], and 7–9 months after the end of the intervention (when the children no longer participated in the lessons) (T3). In both intervention years, the children in the intervention group participated in the F&V lessons for 22 weeks, three times a week; and the children in the control group participated in the regular sedentary classroom lessons. In the first year, six recently graduated and qualified teachers taught the F&V lessons. The regular classroom teachers taught the lessons in the second year (Mullender-Wijnsma et al. 2016). Because the intervention included physically active academic lessons, blinding of children and teachers to group assignment was not possible.

The spelling and math test from the child academic monitoring system were administered by the participating schools at fixed time points. The One-Minute test and the Speed-Test arithmetic were administered by trained test administrators (Mullender-Wijnsma et al. 2016). Follow-up measurements (T3) of the One-Minute test and the Speed-Test were administered 7 months after the completion of the intervention. The participating schools administered the spelling and math test from the child academic monitoring system nine months after the intervention. The One-Minute test and the Speed-Test arithmetic were administered 2 months before the tests from the child academic monitoring system because of practical considerations, such as a school holiday period.

Data Analysis

The Statistical Package for the Social Sciences (version 23.0) was used to calculate the pretest characteristics; the significance level was set at 0.05. Baseline differences between intervention and control group, and between SDC and non-SDC, were examined using an independent t test (age) or a Chi Square Test (grade, gender, SDC).

Repeated measures multilevel modeling (MLwiN 2.29) was used to examine the effects of the F&V intervention at 7–9 months follow-up and to examine the effects of the intervention for SDC. To account for multiple testing, we used Bonferroni correction, resulting in a significance level of 0.0031. Multilevel models were calculated for each academic achievement posttest, with time (T0, T1, T2, T3) as level-one units, children as level-two units, and schools as level-three units. The first model contained the covariates grade, gender, and time. SDC and the interaction between time and SDC were also entered in the first model in order to examine if SDC scored lower on the academic achievement tests.

Subsequently, to assess the effect of the intervention on academic achievement at T3 (aim one), we entered condition and the interaction between condition and time as possible predictors in model 2. The interaction between condition and time was entered to see what happened at the different time points, up and above differences that may be already present at T0. Significant interaction effects would mean that there is an effect of the intervention. To assess the effect of the intervention for SDC at T1, T2, and T3 (aim two), the interaction between condition and SDC and the interaction between condition, time, and SDC were entered (model three). Significant interaction effects would mean that the effect of the intervention differs between SDC and non-SDC. Effect sizes (ES) were calculated as (estimated intervention effect)/√(variance at student level) (Hedges 2007).

Results

Figure 1 shows that 249 children were assigned to the intervention and 250 children to the control group. Two schools dropped out in the second intervention year: one because of the long-term absence of the teacher and the other because it was closed down. At T3 between 322 and 336 children were measured. Common reasons for not completing the tests were absence from school or leaving to attend another school.

The demographic characteristics of the children are shown in Table 1. The proportion of SDC and boys was similar in the control and intervention groups. Because there were more third-grade children in the control group ($\chi^2 = 5.2, p < .05$), the children in the control group were

Fig. 1 Flow of schools and students from enrollment, allocation and analysis

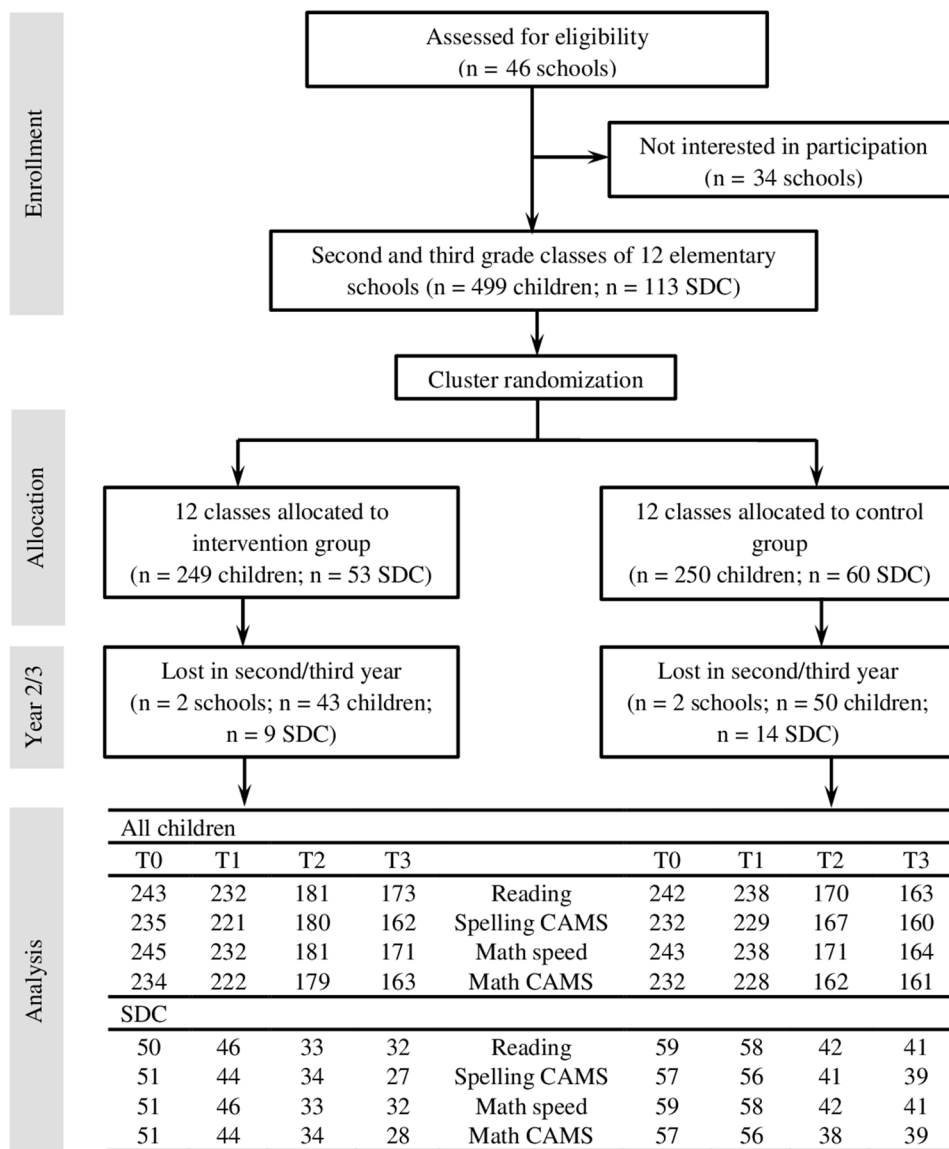


Table 1 Pretest characteristics, by condition and by SDC and non-SDC

| | Control (n=250) | F&V (n=249) | p value | SDC (n=113) | non-SDC (n=386) | p value |
|----------------------|-----------------|-------------|-------------------|-------------|-----------------|-------------------|
| Mean age, years (sd) | 8.15 | 8.01 | 0.03 ^a | 8.26 (0.74) | 8.03 (0.72) | <.01 ^a |
| Second grade [n (%)] | 116 (46.4) | 141 (56.6) | 0.02 ^b | 56 (49.6) | 201 (52.1) | 0.64 ^b |
| Gender, n boys (%) | 110 (44.0) | 116 (46.4) | 0.56 ^b | 54 (47.8) | 172 (44.6) | 0.54 ^b |
| SDC | 60 (24.0) | 53 (21.3) | 0.47 ^b | | | |

^aIndependent *t* test

^bChi square test

Table 2 Mean scores on academic achievement tests with (sd); n

| | T0 | T1 | T2 | T3 |
|-----------------------|--------------------|---------------------|---------------------|---------------------|
| Reading | | | | |
| Intervention group | 80.73 (29.95); 243 | 102.62 (28.89); 232 | 125.34 (26.01); 181 | 138.68 (27.51); 173 |
| SDC | 77.40 (31.79); 50 | 101.28 (31.62); 46 | 121.55 (29.81); 33 | 135.31 (31.79); 32 |
| Non-SDC | 81.59 (29.48); 193 | 102.96 (28.26); 186 | 126.19 (25.12); 148 | 139.44 (26.51); 141 |
| Control group | 80.81 (31.29); 242 | 103.45 (29.05); 238 | 124.82 (30.93); 170 | 134.99 (30.61); 163 |
| SDC | 76.56 (32.34); 59 | 95.36 (26.87); 58 | 115.83 (28.17); 42 | 126.44 (28.82); 41 |
| Non-SDC | 82.17 (30.92); 183 | 106.06 (29.31); 180 | 127.77 (31.33); 128 | 137.87 (30.76); 122 |
| Spelling CAMS* | | | | |
| Intervention group | 117.85 (7.40); 235 | 126.86 (7.90); 221 | 135.12 (8.72); 180 | 137.82 (7.44); 162 |
| SDC | 117.24 (5.74); 51 | 125.73 (8.39); 44 | 134.09 (9.09); 34 | 136.15 (6.40); 27 |
| Non-SDC | 118.02 (7.80); 184 | 127.15 (7.76); 177 | 135.36 (8.64); 146 | 137.32 (7.63); 135 |
| Control group | 118.84 (7.98); 232 | 127.16 (9.09); 229 | 133.49 (9.07); 167 | 136.80 (7.99); 160 |
| SDC | 116.77 (7.58); 57 | 124.88 (8.65); 56 | 130.27 (7.76); 41 | 133.64 (8.67); 39 |
| Non-SDC | 119.52 (8.01); 175 | 127.90 (9.13); 173 | 134.54 (9.24); 126 | 137.82 (7.51); 121 |
| Math speed | | | | |
| Intervention group | 38.82 (17.49); 245 | 60.44 (22.51); 232 | 83.81 (28.16); 181 | 91.92 (26.02); 171 |
| SDC | 36.33 (16.63); 51 | 58.33 (19.22); 46 | 75.64 (24.31); 33 | 85.41 (22.20); 32 |
| Non-SDC | 39.47 (17.70); 194 | 60.97 (23.26); 186 | 85.63 (28.71); 148 | 93.42 (26.67); 139 |
| Control group | 42.05 (18.70); 243 | 61.01 (23.53); 238 | 78.35 (26.59); 171 | 87.65 (26.00); 164 |
| SDC | 42.00 (18.49); 59 | 58.31 (21.04); 58 | 72.43 (21.25); 42 | 83.27 (24.47); 41 |
| Non-SDC | 42.06 (18.82); 184 | 61.88 (24.28); 180 | 80.28 (27.91); 129 | 89.11 (26.41); 123 |
| Math CAMS* | | | | |
| Intervention group | 48.03 (19.16); 234 | 69.58 (16.78); 222 | 82.36 (15.85); 179 | 91.57 (17.58); 163 |
| SDC | 41.82 (18.40); 51 | 62.64 (18.05); 44 | 73.03 (16.89); 34 | 83.82 (16.76); 28 |
| Non-SDC | 49.76 (19.05); 183 | 71.29 (16.04); 178 | 84.54 (14.93); 145 | 93.18 (17.37); 135 |
| Control group | 53.94 (18.14); 232 | 68.65 (17.63); 228 | 82.83 (16.68); 162 | 89.15 (17.95); 161 |
| SDC | 49.53 (17.84); 57 | 65.07 (17.63); 56 | 76.08 (16.79); 38 | 82.95 (17.07); 39 |
| Non-SDC | 55.37 (18.06); 175 | 69.81 (17.53); 172 | 84.90 (16.15); 124 | 91.13 (17.84); 122 |

*CAMS child academic monitoring system

significantly older than the children in the intervention group ($t=2.2, p<.05$). Furthermore, the proportions of boys and children in second grade were similar among SDC and non-SDC, but SDC were significantly older than non-SDC ($t=3.0, p<.05$). Table 2 presents the mean scores and the number of children that took the academic achievement tests per measurement moment.

The first aim of the study was to examine the follow-up effects of the 2-year intervention for all participating

children. The results of the second model of the multilevel analysis can be found in Table 3. For each academic achievement test, except the reading test, ($\Delta\chi^2(4)=3.9, p=0.42$) inserting condition and the interaction between condition and time did significantly improve the first model ($\Delta\chi^2(4)=20.5-50.5, p<0.01$). The results showed that at T0 the score on the math test from the child academic monitoring system of the intervention group was significantly lower than that of the control group ($t=-2.99; p<0.003$).

Table 3 Multilevel regression coefficients (B) and standard error (SE) for each factor predicting the achievement (model 2)

| | Reading | | Spelling (CAMS) | | Math speed | | Math (CAMS) | |
|---------------------------|-----------|--------|-----------------|-------|------------|--------|-------------|--------|
| | B | SE | B | SE | B | SE | B | SE |
| Fixed effects | | | | | | | | |
| Random intercept | 68.713 | 2.400 | 115.236 | 0.695 | 28.969 | 1.950 | 43.990 | 1.429 |
| 3rd grade ^a | 25.139* | 2.291 | 8.284* | 0.570 | 24.216* | 1.591 | 16.892* | 1.227 |
| Boy ^b | 0.467 | 2.289 | -1.279 | 0.572 | 3.223 | 1.594 | 5.213* | 1.226 |
| T1 ^c | 22.093* | 1.000 | 8.539* | 0.409 | 18.860* | 0.987 | 14.963* | 0.793 |
| T2 ^c | 41.908* | 1.135 | 14.786* | 0.459 | 37.497* | 1.115 | 28.356* | 0.895 |
| T3 ^c | 53.665* | 1.155 | 18.126* | 0.467 | 46.556* | 1.135 | 34.493* | 0.904 |
| SDC ^d | -5.412 | 2.864 | -1.831 | 0.776 | -2.565 | 2.123 | -6.908* | 1.629 |
| SDC×T1 | -1.749 | 1.583 | -0.797 | 0.641 | -1.643 | 1.559 | -0.738 | 1.244 |
| SDC×T2 | -3.439 | 1.785 | -1.008 | 0.713 | -6.605* | 1.753 | -2.581 | 1.397 |
| SDC×T3 | -4.176 | 1.802 | -1.526 | 0.742 | -4.527 | 1.775 | -1.180 | 1.436 |
| Intervention ^e | 1.979 | 2.408 | -0.094 | 0.645 | -1.632 | 1.752 | -4.070* | 1.360 |
| Intervention ×T1 | 0.016 | 1.313 | 0.791 | 0.536 | 3.138 | 1.295 | 6.558* | 1.040 |
| Intervention ×T2 | 1.257 | 1.467 | 2.520* | 0.591 | 7.446 | 1.441 | 5.125* | 1.152 |
| Intervention ×T3 | 2.089 | 1.490 | 0.798 | 0.607 | 5.854 | 1.467 | 6.849* | 1.175 |
| Random effects | | | | | | | | |
| Variance schools | 0.000 | 0.000 | 1.032 | 0.816 | 9.990 | 7.168 | 3.076 | 3.077 |
| Variance students | 609.753* | 40.893 | 33.174* | 2.511 | 273.061* | 19.708 | 158.756 | 11.653 |
| Variance time | 99.619* | 4.163 | 15.817* | 0.675 | 97.670* | 4.074 | 59.322 | 2.538 |
| Deviance | 13705.570 | | 9870.660 | | 13347.630 | | 12031.506 | |

^{a,b,c,d,e}respectively 2nd grade, girls, T0, non-SDC, and control group were the reference categories

*Significant ($p < .00313$) regression coefficient

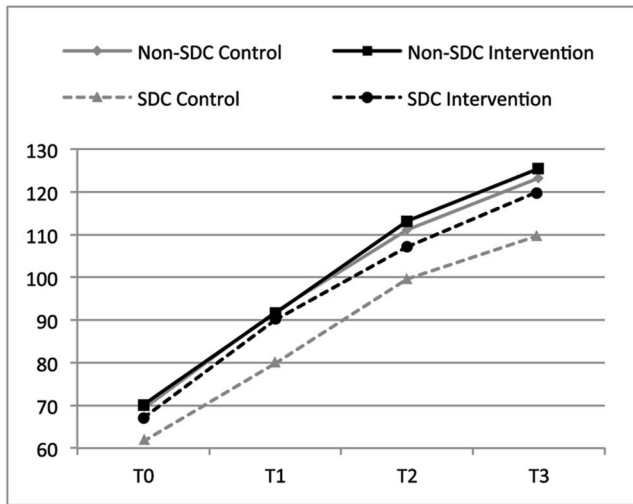
No differences between the intervention and the control group were found on the reading, spelling, and math speed test scores at T0. The results further revealed no significant follow-up effect of the intervention on the reading scores ($t = 1.40$; $p = 0.16$; $ES = 0.08$; Fig. 2a) and spelling scores ($t = 1.31$; $p = 0.19$; $ES = 0.14$; Fig. 2b). The results of the math speed test ($t = 3.99$; $p < 0.005$; $ES = 0.35$; Fig. 2c) and math test from the child academic monitoring system ($t = 5.83$; $p < 0.005$; $ES = 0.54$; Fig. 2d) revealed that the scores of the children in the intervention group had improved significantly more (with respect to T0) than those of the control group at follow-up.

The second aim of the study was to examine the effects of the intervention especially for SDC. SDC was a negative predictor of academic achievement for all achievement tests, indicating that SDC scored lower than non-SDC on the academic achievement tests. For each academic achievement test, inserting the interaction between condition and SDC and the interaction between condition, time, and SDC did not significantly improve the model ($\Delta\chi^2(4) = 2.5\text{--}5.8$, $p > 0.05$). The results revealed no significant different intervention effects between SDC and non-SDC—either after one or after two intervention years or at follow-up—on all four academic achievement tests. This indicates that the intervention did not affect SDC and non-SDC differently.

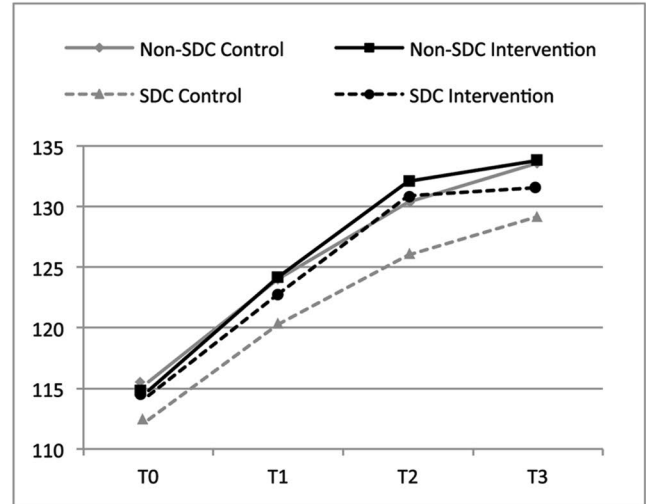
Discussion

In the current study we examined the follow-up effects of the F&V intervention and the effects of the F&V intervention for SDC. It builds on a previous study of F&V lessons, in which effects of a 2-year intervention on math and spelling were found (Mullender-Wijnsma et al. 2016). The results of the current study indicated that effects on math were still apparent at follow-up, but no follow-up effects on spelling were found. The focus of the F&V lessons was on constant practice and repetition (Mullender-Wijnsma et al. 2015b). During the spelling lessons children had to spell words from specific word categories. It is likely that the knowledge acquired during these lessons was not very useful in later (regular) spelling lessons because other word categories were practiced. The mathematical knowledge acquired during F&V lessons might be more useful in later (regular) math lessons because constant practice of math problems during F&V lessons might provide an additional basis for recognition of more difficult math problems. The lasting effects on math indicate that gains in math achievement can be maintained at least 7–9 months after physically active academic lessons end. However, since the effects on spelling disappeared when the F&V lessons were no longer taught, schools may have to encourage their staff to use physically active academic lessons throughout elementary

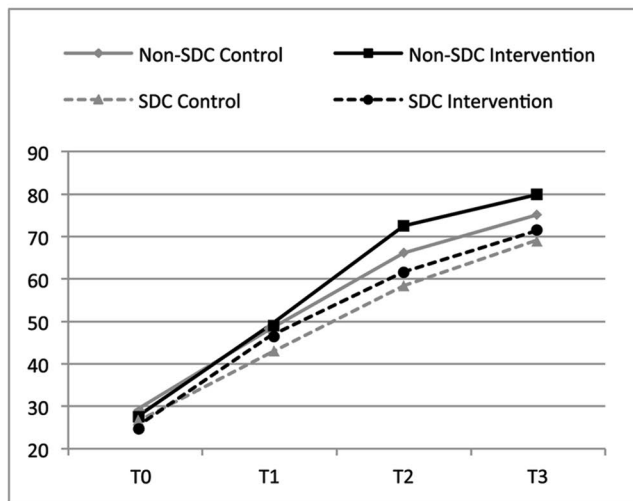
A Reading



B Spelling CAMS



C Math speed



D Math CAMS

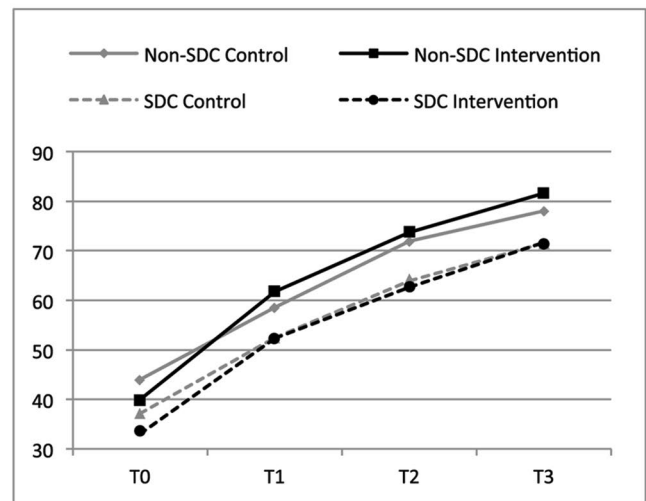


Fig. 2 Predicted mean scores (based on the third model of the multilevel analysis) on the reading (a), spelling test from the child academic monitoring system (b), Math speed (c), and Math test from the child academic monitoring system (d) tests per measurement moment (T0, T1, T2)

school. Donnelly et al. (2009) showed that 95% of the teachers continued to use the physically active academic lessons after the intervention stopped. This is a hopeful finding, but more research is necessary to assess the sustainability of physically active academic lesson programs in elementary schools and to discover the additional learning gains of continued use of the lessons.

The F&V lessons improved the math and spelling performance of children from low-socioeconomic status families (SDC). SDC in the intervention group showed more improvement than SDC in the control group, but SDC did not benefit more from the intervention than non-SDC. These findings correspond with the findings from a previous study on acute effects of F&V lessons (Mullender-Wijnsma et al.

2015a). According to this previous study, the on-task behavior of SDC is generally lower than the on-task behavior of non-SDC. Both SDC and non-SDC benefited from the F&V lessons; their time-on-task was higher immediately after an F&V lesson. However, SDC did not benefit more from the lessons than non-SDC. A study that examined relationships between children’s literacy and numeracy scores and their health behaviors showed that physical activity independently predicted children’s achievement scores after socioeconomic status was controlled for (O’Dea and Mugridge 2012). This suggests that the academic achievement of children from low and high socio-economic status groups is influenced equally by physical activity. Therefore, both SDC as non-SDC can

be provided with physically active academic lessons and benefit academically.

Because time-on-task is positively related to academic achievement (Scheerens et al. 2013), the increased time-on-task immediately after an F&V lesson may be one of the mechanisms that caused the improved academic achievement scores. Other working mechanisms that may have contributed to the effects on children's academic achievement are the effect of physical activity on the brain (Best 2010; Hillman et al. 2008) and the sensorimotor information that the children obtained during the F&V lesson (Kontra et al. 2012; Paas and Sweller 2012). Further research is necessary to examine through which mechanism physically active academic lessons improve academic achievement.

Strengths and Limitations

The current study has some limitations that should be acknowledged. First, administration of the tests from the child academic monitoring system was done by the schools themselves. Teacher expectations may have led to teachers influencing children's test results. However, the teachers knew under what conditions they had to administer a test; therefore it is not likely that they had a large influence on the children's test results. Second, the score on the math test from the child academic monitoring system of the intervention group was significantly lower at baseline. Given that there was more room for improvement, this could possibly have influenced the results. The strengths of this study were the large sample size, the design (cluster-randomized controlled trial), and that long term effects of physically active academic lessons were investigated.

Conclusions

The F&V lessons positively influenced the math and spelling scores of SDC. Both SDC and non-SDC benefited, but SDC did not benefit more from the lessons than non-SDC. Furthermore, at 7–9 months follow-up, the effects of F&V physically active math and language lessons were found to have lasted on math. No follow-up effects were found on spelling and reading. The findings suggest that physically active academic lessons should be considered for inclusion in school curriculums in order to improve the academic achievement of SDC and non-SDC.

Acknowledgements The authors wish to thank the intervention teachers and the participating schools for their cooperation in this study. The language editor of a near-final draft of this paper was P. Goldrick.

Funding This work was supported by the Ministry of Education, Culture and Science (Grant No. ODB10015).

Compliance with Ethical Standards

Ethical Approval All procedures were approved by the ethical committee of the Center for Human Movement Sciences of the University Medical Center Groningen/University of Groningen.

Informed Consent The school principals of the participating schools gave written informed consent. The parents/legal guardians were informed about the F&V project before the start of the intervention. The parents had the option of withdrawing their permission for their child to participate at any time.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- 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.
- Brus, B. T., & Voeten, M. (1973). *One minute test: Justification and Manual*. Nijmegen: Berkhout.
- 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.
- de Vos, T. (1992). *Speed test arithmetic: Manual*. Nijmegen: Berkhout.
- De Vries, S. I., Bakker, I., Van Overbeek, K., Boer, N. D., Hopman-Rock, M., & en Jeugd, P. W. (2005). *Kinderen in Prioriteitswijken: Lichamelijke (In)activiteit en Overgewicht [Children from Priority Neighborhoods: Physical (In)activity and Overweight]*. Leiden: TNO Quality of Life.
- De Wijs, A., Kamphuis, F., Kleintjes, F., & Tomesen, M. (2010). *Spelling for grade 1 to 4: Scientific justification*. Arnhem: Cito.
- 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.
- Donnelly, J. E., Hillman, C. H., Green, J. L., Hansen, D. M., Gibson, C. A., Sullivan, D. K., et al. (2017). Physical activity and academic achievement across the curriculum: Results from a 3-year cluster-randomized trial. *Preventive Medicine, 99*, 140–145.
- Donnelly, J. E., & Lambourne, K. (2011). Classroom-based physical activity, cognition, and academic achievement. *Preventive Medicine, 52*(1), S36–S42.
- Driessen, G., & Dekkers, H. (2008). Dutch policies on socio-economic and ethnic inequality in education. *International Social Science Journal, 59*(193–194), 449–464.
- Greenwood, C. R., Horton, B. T., & Utley, C. A. (2002). Academic engagement: Current perspectives in research and practice. *School Psychology Review, 31*(3), 328–349.
- Grieco, L. A., Jowers, E. M., & Bartholomew, J. B. (2009). Physically active academic lessons and time on task: The moderating effect of body mass index. *Medicine and Science in Sports and Exercise, 41*(10), 1921–1926.

- Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, 32(4), 341–370.
- 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.
- Janssen, J., & Hickendorff, M. (2008). *Categories analysis for the mathematics CAMS test*. Arnhem: Cito.
- Janssen, J., Verhelst, N., Engelen, R., & Scheltens, F. (2010). *Mathematics for grade 1 to 6: Scientific justification*. Arnhem: Cito.
- Kontra, C., Goldin-Meadow, S., & Beilock, S. L. (2012). Embodied learning across the life span. *Topics in Cognitive Science*, 4(4), 731–739.
- Linnenbrink, E. A., & Pintrich, P. R. (2002). Motivation as an enabler for academic success. *School Psychology Review*, 31(3), 313–327.
- Mahar, M. T., Murphy, S. K., Rowe, D. A., Golden, J., Shields, A. T., & Raedeke, T. D. (2006). Effects of a classroom-based program on physical activity and on-task behavior. *Medicine and Science in Sports and Exercise*, 38(12), 2086–2094.
- Ministry of Education, Culture and Science. (2006). *Brochure new weighting system elementary school*. Zoetermeer: CFI.
- Miranda Fredriks, A., Van Buuren, S., Hira Sing, R. A., Maarten Wit, J., & Pauline Verloove-vanhorick, S. (2005). Alarming prevalences of overweight and obesity for children of Turkish, Moroccan and Dutch origin in The Netherlands according to international standards. *Acta Paediatrica*, 94(4), 496–498.
- Mullender-Wijnsma, M. J., Hartman, E., de Greeff, J. W., Bosker, R. J., Doolaard, S., & Visscher, C. (2015a). 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.
- Mullender-Wijnsma, M. J., Hartman, E., de Greeff, J. W., Bosker, R. J., Doolaard, S., & Visscher, C. (2015b). Improving academic performance of school-age children by physical activity in the classroom: 1-year program evaluation. *Journal of School Health*, 85(6), 365–371.
- Mullender-Wijnsma, M. J., Hartman, E., de Greeff, J. W., Doolaard, S., Bosker, R. J., & Visscher, C. (2016). Physically active math and language lessons improve academic achievement: A cluster randomized controlled trial. *Pediatrics*, 137(3), 1–9.
- 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.
- O’dea, J. A., & Mugridge, A. C. (2012). Nutritional quality of breakfast and physical activity independently predict the literacy and numeracy scores of children after adjusting for socioeconomic status. *Health Education Research*, 27(6), 975–985.
- Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review*, 24(1), 27–45.
- Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In R. J. Murnane & G. J. Duncan (Eds.), *Whither opportunity? Rising inequality and the uncertain life chances of low-income children* (pp. 91–116). New York: Russell Sage Foundation.
- Rothstein, R. (2009). Class and the classroom. In K. Ryan & J. M. Cooper (Eds.), *Kaleidoscope: Contemporary and classic readings in education* (pp. 283–289). Belmont, CA: Wadsworth.
- Scheerens, J., Hendriks, M. A., Luyten, J. W., Slegers, P. J. C., & Glas, C. A. (2013). *Productive time in education A review of the effectiveness of teaching time at school, homework and extended time outside school hours*. Den Haag: NWO/PROO.
- Spybrook, J., & Raudenbush, S. W. (2008). *Optimal design software*. Ann Arbor, MI: University of Michigan.
- Vazou, S., Gavrilou, P., Mamalaki, E., Papanastasiou, A., & Sioumala, N. (2012). Does integrating physical activity in the elementary school classroom influence academic motivation? *International Journal of Sport and Exercise Psychology*, 10(4), 251–263.
- Watson, A., Timperio, A., Brown, H., Best, K., & Hesketh, K. D. (2017). Effect of classroom-based physical activity interventions on academic and physical activity outcomes: A systematic review and meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity*, 14(1), 114.
- Webster, C. A., Russ, L., Vazou, S., Goh, T. L., & Erwin, H. (2015). Integrating movement in academic classrooms: Understanding, applying and advancing the knowledge base. *Obesity Reviews*, 16(8), 691–701.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.