Lima, R. A., Larsen, L. R., Bugge, A., Andersen, L. B. (2018). Physical fitness is longitudinally associated with academic performance during childhood and adolescence, and waist circumference mediated the relationship. *Pediatric Exercise Science, 30*, 317-325

Title:

Physical fitness is longitudinally associated with academic performance during childhood and adolescence and waist circumference mediated the relationship

Running Head:

Fitness association with academic performance

Abstract

Purpose: The current investigation aimed to investigate the longitudinal association between physical fitness and academic performance over three years in adolescents. A secondary aim was to determine to what extent waist circumference (WC) mediated the association between physical fitness and academic performance. **Method:** For the current study, 1,020 students from first grade (mean age 7.87 ±0.34 years) to fifth grade (mean age 11.87 ±0.37 years) were monitored annually for three years (2010-2013). Physical fitness was assessed using the Andersen test, 5X5-meter shuttle run, jump height and grip strength tests and by constructing a composite score combining all four fitness tests. Academic performance was assessed by national standardized tests in Danish language and math. Generalized structural equation modeling was used to evaluate the relationships between these variables. **Results:** The Andersen test (standardized β= 0.15 SD), shuttle run (β= -0.18 SD), jump height (β= 0.10 SD) and the fitness composite score (β= 0.23 SD) were positively associated with academic performance over three years. In addition, WC partially mediated the association between physical fitness and academic performance. **Conclusion:** Thus, physical fitness abilities should be stimulated during childhood and early adolescence because of their positive association with academic performance.

Keywords: aerobic; adolescent; physical fitness; body composition; strength.

Introduction

There is a constant pressure of the society on children and adolescents to present higher academic performance and this has led to a decrease in the amount of time students are exposed to physical activities at school in several countries (23). Concerns about the possible deleterious impact of physical inactivity in schools on student's health and academic performance encouraged the research community to increase the focus on the association between physical activity, physical fitness and academic performance in children and adolescents (3). It has been shown that higher physical fitness level, especially cardiorespiratory fitness, is positively associated with better academic performance in children and adolescents (3,18). However, most of the evidence is based on cross-sectional studies (3,12,18,21,25).

Two recent systematic reviews summarized the evidence of the studies evaluating the association between physical fitness and academic performance (3,18). In the review by Donnelly, et al. (2016) all three longitudinal studies included, reported a positive relationship between physical fitness and academic performance. However, Santana, et al. (2017) only found positive longitudinal association between physical fitness and academic performance in four out of the seven studies included in their review. Moreover, only cardiorespiratory fitness was assessed in the majority of the studies in both reviews, limiting the external validity of the findings to include only this aspect of physical fitness (3,18). In addition, it is not clear whether the association between physical fitness and academic performance differs for boys and girls (10,12,15,18).

Although excess body mass has been negatively associated with physical fitness (13,17) and academic performance (12,21), the associations between physical fitness, excess body mass and academic performance have not been fully elucidated. The available evidence suggests that physical fitness has a stronger association with academic performance than excess body mass (12,21) and that participants with high body fatness presented higher academic performance if they were fit (5), but it has not been tested whether excess body mass mediates part of the relationship between physical fitness and academic performance (3,18). Therefore, the aim of this study was to investigate the association between physical fitness and academic performance in a three-year longitudinal study. A secondary aim was to determine to what extent waist circumference mediated the association between physical fitness and academic performance.

Participants and Methods

The current investigation was part of the CHAMPS-study DK, which was the research part of a community initiative: the Svendborg Project. The Svendborg Project started in 2008 and all 19 public schools in the municipality of Svendborg, Denmark, were invited to participate. Six schools agreed to be enrolled into the project as intervention schools and four demographically similar schools were assigned as control schools, also from the municipality of Svendborg. The study was quasi-experimental. The program included students from kindergarten to fourth grade in 2008. The students at the intervention schools received a minimum of 4.5 hours of physical education (PE) per week divided over at least three sessions of at least 60 minutes each. To ensure optimal quality of the PE lessons, PE teachers and pedagogues participated in a 40-lesson course in "Age-related concepts of training" organized by Team Denmark in collaboration with the faculty of education at the University College Lillebaelt. The goal was to ensure targeted and responsible training for the growing individual (16). The control schools received their regular PE program (1.5 hours/week). The overall aim of the CHAMPS-study DK was to determine the health benefits of increased PA through age-related training for the participants (24). The study was conducted in accordance with the Declaration of Helsinki. It was approved by the local scientific ethics committee (ID S20080047) and registered in the Danish Data Protection Agency (J.nr. 2008-41-2240). Parents had to provide written consent before their children were included in the study. Because the complete methodology of the CHAMPS-study DK had been published elsewhere (24), only the variables of interest are described here.

Measurements

In 2010, Denmark started a national standardized test system to assess academic performance of all students enrolled in Danish schools. Because no data was available for academic performance from 2008-9, the current study evaluated the association between physical fitness and academic performance in 918 participants during a three-year period (2010 – 2013). The academic tests were computerized standardized tests, which progressed according to the students' performance. For instance, if the student answered a question correct, the next question became more complex. Similarly, if the student answered a question incorrect, the next question became less complex. The grading was standardized, ranging from 0 to 300 points; therefore, a score of 50 in math was comparable to a score of 50 in Danish language. Participants were tested in Danish language in second, forth, sixth and eighth grades, and in math in third and sixth grades. Additionally, both Danish language and math were evaluated in three different domains (scores ranged from 0 to 100 for each of them). In particular, language understanding, decoding and text comprehension were domains evaluated in the Danish language tests, and; algebra, geometry and basic math skills were domains tested in math. In addition to the Danish language and math tests, a composite academic score was made combining all tests.

Information on the participant's mother's education was collected via a questionnaire and categorized into five levels: primary and lower secondary education, general upper secondary education, vocational education and training, bachelor degree, and master or PhD degree. From 2010 to 2013, during spring, all participants were physically examined and tested once a year. Height was measured to the nearest 0.5 centimeter (cm) with a portable stadiometer (Seca 214, Seca Corporation, Hanover, MD). WC was measured with a measuring tape at umbilicus level to the nearest 0.5 cm at the end of a gentle expiration. Two measurements were performed and in case of a difference between the measurements greater than one cm, a third measurement was performed. The average of the two nearest measurements was calculated and used for the analysis. For the current study, WC is the proxy measure of excess body mass.

Physical fitness was measured by four different tests. The Andersen test is an intermittend running test lasting 10 minutes and the distance covered was used a measure of cardiorespiratory fitness according to the test protocol (1). The shuttle run was a test of agility and cardiorespiratory capacity, which consists of five laps on a five-meter lane, completed as fast as possible. It was scores as number of seconds and a low number was a better score (22). The vertical jump test measured the highest jump in centimeters after a minimum of three trials, but the participant had up to six trials if they kept improving their jump. This protocol followed that of the Abalakow vertical jump test (9), and was considered a proxy measure of functional dynamic strength in the lower extremity. Hand grip strength was measured to test the isometric strength of the arms (11). The better of two trials was recorded in kilograms using a dynamometer (JAMAR dynamometer, Scandidact, Cat.No.281128).

Statistics

STATA version 14.0 (StataCorp LP, College Station, TX, USA) was used for all analyses. For descriptive purposes, means and standard deviations are presented and ANOVA was performed to test for differences between boys and girls and school grades. Generalized structural equation modeling (SEM) was used to estimate whether: i) physical fitness (any of the fitness tests or fitness composite score) longitudinally were associated with academic performance (Danish language, math, and the composite academic score), ii) WC mediated the association between physical fitness and academic performance, and iii) the difference in academic performance between different tertiles of physical fitness.

The composite fitness score was constructed as the sum of z-scores of the Andersen test, inverse of 5X5-m shuttle run, grip strength and jump height. The composite score for academic performance was the sum of the three Danish language scores and the three math scores.

Supplemental material 1 (Figure 1) presents the pathways examined in the analysis. Note that all follow-up measures were included in one box for each respective variable in Figure 1 because it is difficult to show the complexity of all variables assessed at all time points within one figure. In summary, the same diagram was developed for each of the fitness tests (Andersen test, 5X5-m shuttle run, jump height, grip strength), and the fitness composite score individually to examine the longitudinal association between physical fitness and academic performance (Danish language, math, and the composite academic score). In addition, WC was examined as a mediating variable for the possible association between physical fitness (fitness tests or fitness latent score) and academic performance (Danish language, math, and the composite academic score). Similarly, the same diagram was used to evaluate whether academic performance differ between different tertiles of physical fitness. All models were controlled for sex, age, school type (intervention and control), class, mother's educational level and height. It was also considered the hierarchical structure of data (repeated observations for each student nested in classes and classes nested in schools). We calculated the variance related to the clusters (school and classes within school) and the intraclass correlation coefficient (ICC) for each model to interpret the variation between school classes, classes, and individuals. There was no interaction between physical fitness and school type (intervention and control) in the longitudinal association between physical fitness and academic performance, and all participants were therefore analyzed together.

Results

Table 1 presents means and standard deviations of the physical fitness tests, waist circumference and academic performance in Danish language and math for the different grades at the first measurement in 2010. As expected, older participants (higher grades) exhibited significantly better performance on the physical fitness tests and higher WC compared to younger participants when analyzing boys and girls together or independently. However, there was no clear pattern of difference in academic performance at the different grades. Girls exhibited higher scores in Danish language than boys, and no sex differences were observed for math scores.

Longitudinal association between physical fitness and waist circumference and between waist circumference and academic performance

Higher WC was negatively associated with academic performance over time, independent of the physical fitness variables entered into the model. In addition, all physical fitness tests (except grip strength) and the composite fitness score were associated with WC. For instance, for each additional standard deviation performed in the Andersen test, WC decreased by 0.362 SD; for each SD in the shuttle run test, WC increased by 0.311 SD. Similarly, WC decreased 0.216 SD for each additional standard deviation jumped in the jump height test and for each standard deviation increase in the fitness composite score, WC decreased 0.301 SD (see Figure 2).

Longitudinal association between physical fitness and composite academic performance

During the three-year follow-up most fitness tests (Andersen test, shuttle run and jump height) and the fitness latent score were longitudinally associated with the composite academic performance score in boys and girls when analyzed separately and together (Table 2). Grip strength was associated with the composite academic performance score for girls, but not for boys. More specifically, in all participants, for each additional SD performed in the Andersen test, the composite academic performance score increased 0.13 SD; and for each 1 SD increase in the shuttle run test, the score decreased 0.13 SD. Moreover, each 1 SD increase in the vertical jump test resulted in a 0.10 SD increase in academic score; and for each 1 SD increase in the composite fitness score, the academic score increased by 0.17 SD (see Table 2).

Longitudinal association between physical fitness and Danish language

Analyzing boys and girls together, the Andersen test, 5X5-m shuttle run, jump height and the composite fitness score were longitudinally associated with Danish language scores directly and also mediated by WC. Comparing boys and girls, just the Andersen test and jump height differed on the association with Danish language performance over time. For boys, the Andersen test was associated with Danish language performance through mediation of WC, while for girls, the Andersen test was associated with Danish language performance both directly as well as through mediation by WC. Alternatively, jump height was associated both directly and mediated by WC with Danish language performance for boys. For girls, jump height was associated with Danish language performance through mediation of WC (data not shown).

Longitudinal association between physical fitness and math performance

For all participants, the Andersen test, 5X5-m shuttle run, jump height and the composite fitness score were longitudinally associated with math performance. In addition, 5X5-m shuttle run and jump height were also associated with math performance via mediation of WC. Furthermore, for boys, none of the physical fitness tests (Andersen test, 5X5-m shuttle run, jump height and grip strength) or the composite fitness score were associated with math performance via mediation by WC. Only grip strength was not associated directly with math performance. For girls, the Andersen test, 5X5-m shuttle run, jump height and the composite fitness score were directly associated and mediated via WC with math performance, while grip strength was associated directly with math performance (data not shown).

Difference in academic performance depending on the physical fitness tertile

Specifically, the Andersen test, the shuttle run test, and the fitness latent score (Fitness score) presented similar increments on academic performance from the lowest to the second (β= 12.36-13.41) and third tertiles (β= 25.06-26.31). In other words, children in the highest physical fitness tertile exhibited around 15% higher academic performance compared to participants in the lowest physical fitness tertile during the three-year follow-up. The increment on academic performance was smaller for vertical jump height (Second tertile: β= 6.78; Third tertile: β= 13.84). For grip strength, no significant difference was observed in academic performance between children in higher grip strength tertiles and participants in the lowest tertile (see Figure 3).

Independent of the physical fitness variable used, the diagram used in the analyses fit the data (Comparative Fit Index – CFI = 0.978-0.991; Tucker Lewis Index – TLI = 0.942-0.973; Root Mean Square Error of Approximation – RMSEA = 0.056-0.065; Standardized Root Mean Residual – SRMR = 0.015-0.024; Coefficient determination – R^2 = 0.699-0.757) – data not shown. We also calculated the Akaike Information Criterion (AIC) for each of the models. AIC in the Andersen test model: 116431.942; AIC in the Shuttle run model: 108084.528; AIC in the Vertical jump height model: 115599.030; AIC in the Grip strength model: 85004.124; AIC in the Fitness score model: 128376.843 – data not shown. Despite the physical fitness variable in the model, most of the variance was in the individual level (Andersen test, Vertical jump height and Fitness score models: School cluster ICC: 4%, Class cluster ICC: 9%, Individual cluster ICC: 56%; Shuttle run model: School cluster ICC: 4%, Class cluster ICC: 8%, Individual cluster ICC: 56%; Grip strength: School cluster ICC: 5%, Class cluster ICC: 10%, Individual cluster ICC: 58%) – data not shown.

Discussion

The current study investigated the longitudinal association between physical fitness and academic performance over three years in children aged 7-13 years at baseline. Four different physical fitness tests and a composite fitness score were used to represent the physical fitness level of the participants, and; additionally, scores in Danish language, math and the composite academic score were the outcomes used to express academic performance. Analyses showed that higher physical fitness was longitudinally associated with better academic performance during childhood, 7-13 years of age, and early adolescence, 10-16 years of age. In addition, waist circumference was negatively associated with academic performance and mediated some of the longitudinal relationship between physical fitness and academic performance.

The Andersen test, 5X5-m shuttle run, jump height and the composite fitness score were longitudinally associated with better academic performance in Danish language and math individually and combined. The findings support previous studies (4,6,12,21,25). Two recent systematic reviews observed a consistent positive longitudinal relationship between physical fitness and academic performance (3,18). Nevertheless, the majority of the longitudinal studies have used cardiorespiratory fitness as their only physical fitness outcome (6,19,21,25). The current study adds to previous findings by demonstrating that agility, speed and jump height were associated with better academic performance over time. Therefore, whole body activities improving cardiorespiratory fitness, agility and lower body strength should be emphasized during childhood. In this age group, grip strength is mainly a fitness component of a small muscle group, which may explain why it was not consistently associated with academic performance.

Importantly, in the present study participants with higher physical fitness (in the highest tertiles) did not present lower academic performance compared to participants with lower physical fitness levels. On the contrary, those with higher physical fitness demonstrated higher academic performance than participants with the lowest physical fitness. Haapala, et al. (2014) observed similar result. Concerns about increased physical activity should have a negative effect on academic performance, have led to a reduction in timetabled physical education in school in several countries, because physical education lessons replaced academic lessons (7). Although improving academic achievement is not the main focus of promoting physical activity, our results suggest that spending time in physical activity should be encouraged at school also as a mean to improve physical fitness (14), because higher physical fitness level is related to better academic performance (6).

High WC was longitudinally associated with poorer academic performance, which has also been observed in previous studies (5,8,12,21,25). Interestingly, waist circumference mediated to some extent the association between physical fitness and academic performance which has been demonstrated for the first time in a longitudinal setting. In other words, the reported results showed that increments in physical fitness during the follow-up consequently decreased participants' waist circumference and this decrease was associated with better academic performance. The observational nature of the present study makes it impossible to infer on causality and the underlying biological mechanisms that may explain the mediation component of waist circumference in the association between physical fitness and academic performance. However, several recent studies have proposed plausible mechanisms for the association between academic performance and body composition, such as evidence showing that excess adiposity might impair cognitive function and thereby academic achievement, and that overweight and obese students tend to have lower IQ (20,26). It is possible that some of these negative consequences of adiposity influence the academic achievement-physical fitness relationship.

The current investigation has some limitations that should be taken into consideration in the interpretation of the results. The CHAMPS-study DK is an intervention initiative with increased number of PE lessons at the intervention schools. Thus, it is possible that the results are somehow influenced by the intervention; however, this concern was considered in the analysis, which was adjusted for school type (intervention/control) and we found no interaction. In addition, no effect of the intervention was found on academic performance (2). Students were not tested for Danish language and math every year, so the amount of information collected for each subject could have been more extensive, but the tests followed the national test system in Denmark. Nevertheless, both Danish language and math were assessed in three different domains, which improved the estimation of the students' academic performance.

Another strength is the long-term follow-up. This investigation was able to follow participants over three years while monitoring different physical fitness components and comprehensive academic performance outcomes. Moreover, only the Andersen test and the 5X5-m shuttle run were possibly collinear ($r<0.7$), which implies that different physical fitness abilities are independently longitudinally associated with academic performance. Furthermore, it is one of the strengths of the current investigation that socioeconomic status (maternal education) was considered in the analysis, which is an improvement compared to several previous investigations (3), and additionally, the current study evaluated the mediation contribution of waist circumference on the association between physical fitness and academic performance.

Conclusions

In conclusion, physical fitness was positively associated with academic performance during childhood and early adolescence, and waist circumference mediated part of this association. Physical activities capable of improving physical fitness should be encouraged both inside and outside school hours for all students.

Acknowledgments

The current study was funded by: TrygFonden; CAPES Scholarship, number: 1361/13-9; The Danish Heart Foundation; The Nordea Foundation; The IMK Foundation; The Region of Southern Denmark; The Egmont Foundation; The A.J. Andersen Foundation; The Danish Rheumatism Association; Østifternes Foundation; Brd. Hartmanns Foundation; TEAM Denmark; University College Lillebaelt Department of Physiotherapy; University of Southern Denmark; The Danish Chiropractic Research Foundation; and The Municipality of Svendborg.

References

1. Andersen LB, Andersen TE, Andersen E, Anderssen SA. An intermittent running test to estimate maximal oxygen uptake: the Andersen test. *J Sports Med Phys Fitness*. 2008;48(4):434–437.

2. Bugge A, Möller S, Tarp J, et al. Influence of a 2- to 6-year physical education intervention on scholastic performance: The CHAMPS study-DK. 2017 May 22 Epub.

3. Donnelly JE, Hillman CH, Castelli D, et al. Physical Activity, Fitness, Cognitive Function, and Academic Achievement in Children: a systematic review. *Med Sci Sport Exerc*. 2016;48(6):1197–1222.

4. Fedewa AL, Ahn S. The effects of physical activity and physical fitness on children’s achievement and cognitive outcomes: a meta-analysis. *Res Q Exerc Sport*. 2011;82(3):521–535.

5. García-Hermoso A, Esteban-Cornejo I, Olloquequi J, Ramírez-Vélez R. Cardiorespiratory Fitness and Muscular Strength as Mediators of the Influence of Fatness on Academic Achievement. 2017 August Epub.

6. Haapala EA, Poikkeus A-M, Tompuri T, et al. Associations of motor and cardiovascular

performance with academic skills in children. *Med Sci Sports Exerc*. 2014;46(5):1016–1024.

7. Hills AP, Dengel DR, Lubans DR. Supporting Public Health Priorities: Recommendations for Physical Education and Physical Activity Promotion in Schools. *Prog Cardiovasc Dis*. 2015;57(4):368–374.

8. Kamijo K, Khan NA, Pontifex MB, et al. The relation of adiposity to cognitive control and scholastic achievement in preadolescent children. *Obesity (Silver Spring)*. 2012;20(12):2406–2411.

9. Klavora P. Vertical-jump Tests: A Critical Review. *Strength Cond J*. 2000;22(5):70.

10. Kwak L, Kremers SPJ, Bergman P, Ruiz JR, Rizzo NS, Sjöström M. Associations between Physical Activity, Fitness, and Academic Achievement. *J Pediatr*. 2009;155(6):914–918.e1.

11. Larsen LR, Kristensen PL, Junge T, Rexen CT, Wedderkopp N. Motor Performance as Predictor of Physical Activity in Children. *Med Sci Sport Exerc*. 2015;47(9):1849–1856.

12. London RA, Castrechini S. A Longitudinal Examination of the Link Between Youth Physical Fitness and Academic Achievement. *J Sch Health*. 2011;81(7):400–408.

13. Lopes VP, Maia JAR, Rodrigues LP, Malina R. Motor coordination, physical activity and fitness as predictors of longitudinal change in adiposity during childhood. *Eur J Sport Sci*. 2012;12(4):384–391.

14. Minatto G, Barbosa Filho VC, Berria J, Petroski EL. School-Based Interventions to Improve Cardiorespiratory Fitness in Adolescents: Systematic Review with Meta-analysis. *Sports Med*. 2016;46(9):1273–1292.

15. Morita N, Nakajima T, Okita K, Ishihara T, Sagawa M, Yamatsu K. Relationships among

fitness, obesity, screen time and academic achievement in Japanese adolescents. 2016 September 1 Epub.

16. Rexen CT, Ersbøll AK, Møller NC, Klakk H, Wedderkopp N, Andersen LB. Effects of extra school-based physical education on overall physical fitness development - the CHAMPS study DK. *Scand J Med Sci Sports*. 2015;25(5):706–715.

17. Rodrigues LP, Stodden DF, Lopes VP. Developmental pathways of change in fitness and motor competence are related to overweight and obesity status at the end of primary school. *J Sci Med Sport*. 2016;19(1):87–92.

18. Santana CCA, Azevedo LB, Cattuzzo MT, Hill JO, Andrade LP, Prado WL. Physical fitness and academic performance in youth: A systematic review. *Scand J Med Sci Sports*. 2017;27(6):579–603.

19. Sardinha LB, Marques A, Minderico C, et al. Longitudinal Relationship between Cardiorespiratory Fitness and Academic Achievement. *Med Sci Sports Exerc*. 2016;48(5):839–844.

20. Smith E, Hay P, Campbell L, Trollor JN. A review of the association between obesity and cognitive function across the lifespan: implications for novel approaches to prevention and treatment. *Obes Rev*. 2011;12(9):740–755.

21. Telford RD, Cunningham RB, Telford RM, Abhayaratna WP, Abharatna WP. Schools with fitter children achieve better literacy and numeracy results: evidence of a school cultural effect. *Pediatr Exerc Sci*. 2012;24(1):45–57.

22. Tomkinson GR, Lang JJ, Tremblay MS, et al. International normative 20 m shuttle run values from 1 142 026 children and youth representing 50 countries. 2016 May 20 Epub.

23. UNESCO. World-wide survey of school physical education. Report 2013. Paris: United Nations Educational, Scientific and Cultural Organization. 2014 Epub.

24. Wedderkopp N, Jespersen E, Franz C, et al. Study protocol. The Childhood Health, Activity, and Motor Performance School Study Denmark (The CHAMPS-study DK). *BMC Pediatr*. 2012;12(1):672.

25. Wittberg RA, Northrup KL, Cottrell LA. Children's aerobic fitness and academic achievement: a longitudinal examination of students during their fifth and seventh grade years. *Am J Public Health*. 2012;102(12):2303–2307.

26. Yu ZB, Han SP, Cao XG, Guo XR. Intelligence in relation to obesity: a systematic review and meta-analysis. *Obes Rev*. 2010;11(9):656–670.

Table 1. Descriptive characteristics presented in mean and standard deviation (SD) of the age, physical fitness tests, weight, height, waist circumference (WC) and academic achievement scores for the different grades at baseline in 2010.

All participants	1st grade (n= 184 – 190)		2nd grade (n= 215 - 225)		3rd grade (n= 220 - 237)		4th grade (n= 217 - 238)		5th grade (n= 210 - 225)	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Age (years)	7.87	(0.34)	8.89	(0.33)	9.86	(0.39)	10.88	(0.32)	11.87	(0.37)
Andersen test (m)	895.30	(80.06)	926.72	(89.58)	958.82	(107.71)	991.66	(101.82)	1007.18	(96.21)
Shuttle run (s)	25.33	(2.57)	24.36	(2.27)	23.16	(2.79)	21.93	(1.92)	21.70	(2.40)
Jump height (cm)	26.23	(5.20)	28.36	(5.39)	29.84	(5.34)	32.10	(5.27)	33.29	(5.92)
Grip strength (kg)	13.38	(2.60)	15.54	(3.24)	17.62	(3.94)	20.35	(4.04)	23.54	(5.38)
WC (cm)	57.56	(5.32)	61.03	(6.22)	63.67	(6.90)	65.27	(7.17)	68.55	(8.27)
Weight (kg)	27.05	(4.23)	30.96	(5.46)	34.13	(6.27)	37.83	(6.85)	42.46	(8.82)
Height (cm)	129.35	(5.55)	135.72	(5.95)	140.61	(6.79)	146.27	(6.42)	152.26	(7.33)
Danish (points)	-	-	162.46	(75.53)	-	-	161.78	(75.39)	-	-
Math (points)	-	-	-	-	166.68	(66.52)	-	-	-	-
Boys	1st grade		2nd grade		3rd grade		4th grade		5th grade	

Age (years)	7.90	(0.36)	8.96	(0.33)	9.94	(0.41)	10.95	(0.32)	11.94	(0.37)
Andersen test (m)	918.73	(83.15)	961.80	(86.10)	998.93	(108.18)	1039.15	(96.61)	1034.88	(106.69)
Shuttle run (s)	24.52	(2.52)	23.76	(2.17)	22.39	(2.34)	21.31	(1.71)	21.41	(2.92)
Jump height (cm)	26.13	(5.15)	28.45	(5.72)	30.80	(5.78)	33.18	(5.53)	34.48	(5.69)
Grip strength (kg)	14.19	(2.76)	16.38	(3.33)	18.97	(3.85)	21.68	(3.87)	24.55	(5.46)
WC (cm)	57.38	(5.68)	61.24	(5.40)	63.68	(6.53)	65.50	(7.14)	68.93	(8.48)
Weight (kg)	27.32	(4.28)	31.21	(4.94)	34.55	(6.07)	37.99	(6.93)	42.81	(9.33)
Height (cm)	130.58	(5.76)	136.40	(5.74)	141.88	(6.97)	146.94	(6.56)	153.21	(7.34)
Danish (points)	-	-	166.94	(75.88)	-	-	154.82	(76.70)	-	-
Math (points)	-	-	-	-	168.75	(71.69)	-	-	-	-
Girls	1st grade		2nd grade		3rd grade		4th grade		5th grade	
Age (years)	7.84	(0.33)	8.84	(0.32)	9.78	(0.35)	10.82	(0.32)	11.81	(0.35)
Andersen test (m)	876.46	(72.57)	897.34	(81.83)	915.68	(89.31)	953.99	(89.65)	981.51	(77.40)
Shuttle run (s)	25.98	(2.42)	24.89	(2.24)	23.98	(3.00)	22.40	(1.94)	21.98	(1.73)
Jump height (cm)	26.31	(5.26)	28.28	(5.12)	28.79	(4.61)	31.26	(4.91)	32.18	(5.94)

Grip strength (kg)	12.72	(2.26)	14.82	(2.99)	16.15	(3.51)	19.32	(3.88)	22.61	(5.16)
WC (cm)	57.71	(5.03)	60.84	(6.85)	63.67	(7.31)	65.10	(7.22)	68.20	(8.10)
Weight (kg)	26.82	(4.19)	30.73	(5.88)	33.66	(6.48)	37.71	(6.81)	42.13	(8.35)
Height (cm)	128.33	(5.17)	135.14	(6.09)	139.21	(6.32)	145.75	(6.28)	151.38	(7.25)
Danish (points)	-	-	170.02	(70.03)	-	-	175.97	(72.28)	-	-
Math (points)	-	-	-	-	164.94	(58.83)	-	-	-	-

Table 2. Parameter estimates of the slope of the total and mediated longitudinal association of the fitness tests and the fitness latent score (Fitness) on the composite academic achievement scores for all participants and stratified by sex.

	Academic achievement					
All participants	Mediated by WC			Total impact		
	Stand. β	β	(95% CI)	Stand. B	β	(95% CI)
Andersen test (m)	0.04	0.03	(0.016: 0.039)	0.15	0.10	(0.069: 0.128)
Shuttle run (s)	-0.03	-1.05	(-1.484: -0.610)	-0.18	-5.76	(-7.101: -4.417)
Jump height (m)	0.03	0.30	(0.200: 0.404)	0.10	1.05	(0.611: 1.487)
Grip strength (kg)	-0.01	-0.05	(-0.151: 0.060)	0.04	0.40	(-0.256: 1.053)
Fitness latent score (z-scores)	0.04	0.04	(0.022: 0.060)	0.23	0.23	(0.182: 0.281)
Boys	Stand. β	β	(95% CI)	Stand. B	β	(95% CI)
Andersen test (m)	0.03	0.02	(0.003: 0.034)	0.12	0.09	(0.041: 0.131)
Shuttle run (s)	-0.02	-0.68	(-1.235: -0.115)	-0.15	-5.08	(-7.146: -3.008)
Jump height (m)	0.02	0.19	(0.072: 0.315)	0.10	1.00	(0.386: 1.614)
Grip strength (kg)	< -0.01	-0.03	(-0.119: 0.059)	-0.03	-0.27	(-1.157: 0.617)
Fitness latent score (z-scores)	0.03	0.03	(0.003: 0.054)	0.20	0.20	(0.121: 0.273)
Girls	Stand. β	β	(95% CI)	Stand. B	β	(95% CI)
Andersen test (m)	0.05	0.03	(0.019: 0.051)	0.17	0.11	(0.075: 0.154)
Shuttle run (s)	-0.05	-1.39	(-2.042: -0.740)	-0.21	-6.38	(-8.131: -4.627)
Jump height (m)	0.04	0.44	(0.269: 0.617)	0.10	1.20	(0.558: 1.847)
Grip strength (kg)	-0.01	-0.10	(-0.317: 0.117)	0.12	1.25	(0.238: 2.255)
Fitness latent score (SD)	0.05	0.05	(0.023: 0.080)	0.27	0.27	(0.201: 0.334)

Note: Sex, age, school type (intervention and control), class, mother's educational level and height controlled all the analyses.

Legend: (m) refers to meters; (s) refers to seconds; (kg) refers to kilograms; (SD) refers to standard deviations; Stand. β refers to standardized Beta coefficients between the exposures and academic performance.

Figure 1

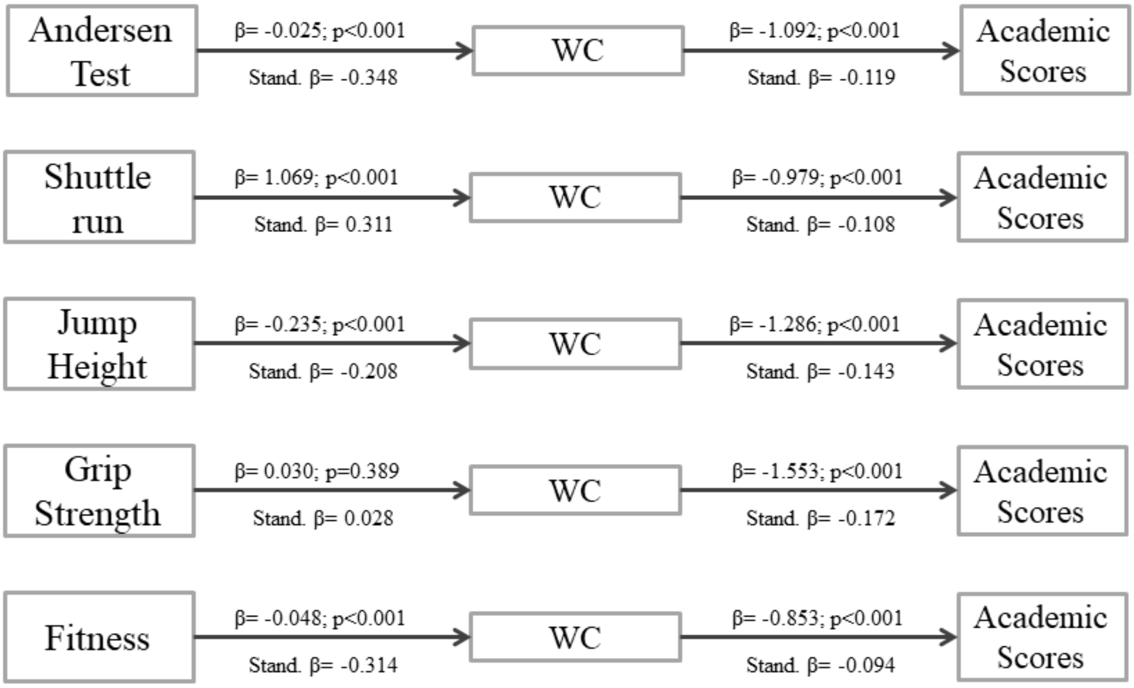


Figure 1. Parameter estimates of the slopes of the longitudinal influence of physical fitness over time on waist circumference (WC) and the influence over time of WC on academic performance scores.

Note: Sex, age, school type (intervention and control), class, mother's educational level and height controlled all the analyses. Fitness refers to fitness latent score. Stand. refers to standardized.

Figure 2

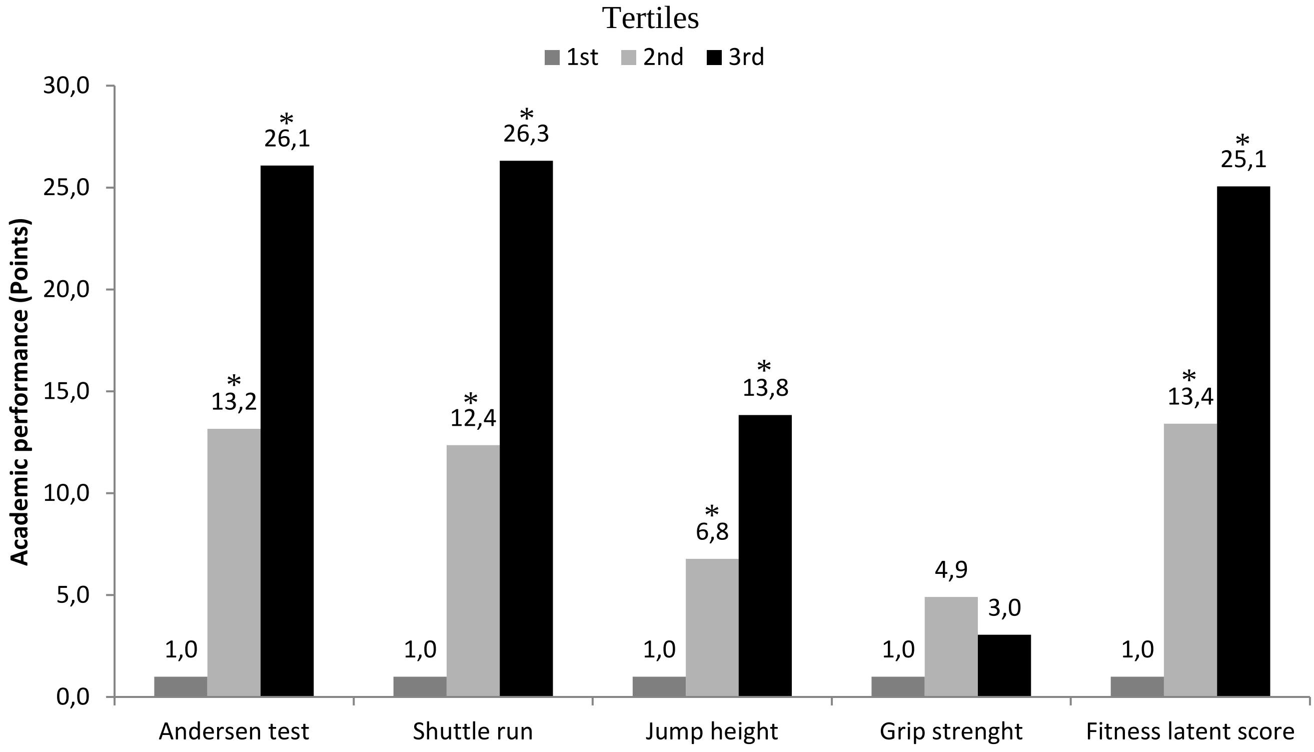
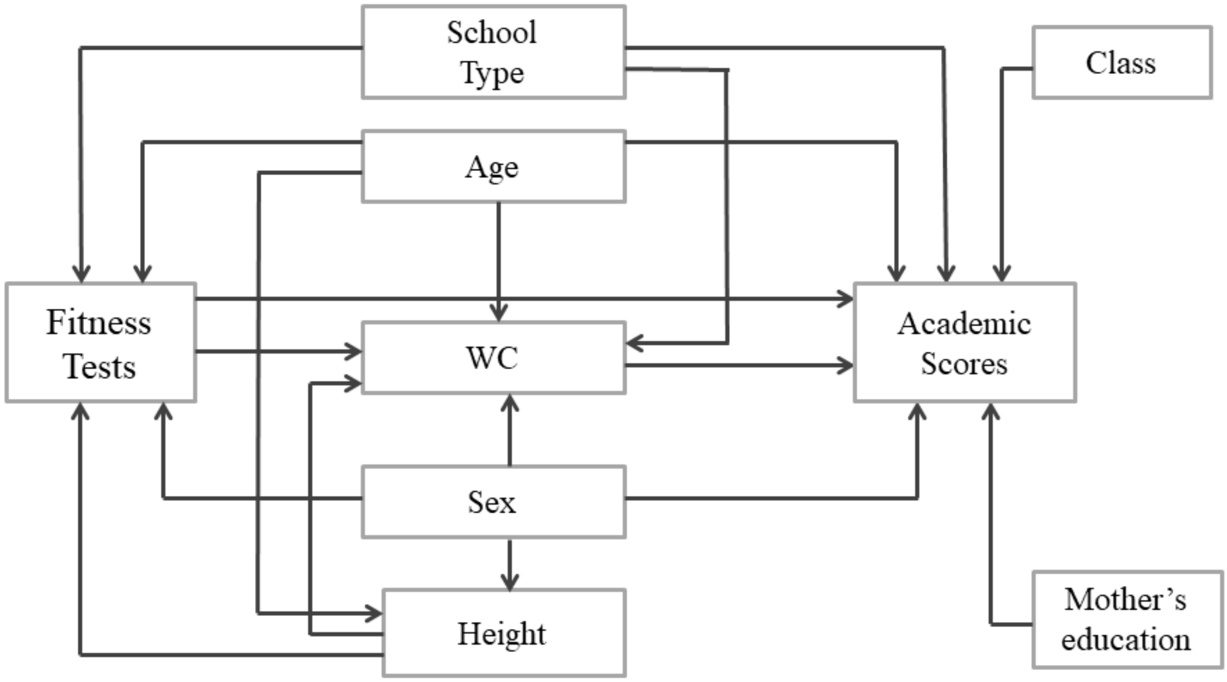


Figure 2. Longitudinal differences in composite academic score between tertiles of physical fitness (*p<0.05 differences between second and third tertiles compared to the first tertile).

Note: Sex, age, school type (intervention and control), class, mother's educational level and height controlled all the analyses.

Supplement 1



Supplemental material 1. Diagram with the pathways used in the structured equation modeling analyses.
Note: Fitness tests is a general term used to refer to the Andersen test, Shuttle run, Vertical jump height, Grip strength, or Fitness latent score. WC refers to Waist circumference. School type refers to assignment to control or intervention school physical education program. Academic scores refers to the grades in Danish, Math, and combined academic score.

Supplemental material 2. Parameter estimates of the slope of the total and mediated influence of the fitness tests and the fitness latent score (Fitness) on Danish language scores for all participants and stratified by sex.

All participants	Danish language scores			
	Mediated by WC		Total impact	
	β	(95% CI)	β	(95% CI)
Andersen test (m)	0.03	(0.019: 0.047)	0.07	(0.037: 0.111)
Shuttle run (s)	-1.16	(-1.699: -0.620)	-5.11	(-6.836: -3.381)
Jump height (m)	0.31	(0.181: 0.432)	0.66	(0.122: 1.207)
Grip strength (kg)	-0.10	(-0.235: 0.034)	0.17	(-0.641: 0.985)
Fitness latent score (z-scores)	0.05	(0.025: 0.074)	0.19	(0.127:0.256)
Boys	β	(95% CI)	β	(95% CI)
Andersen test (m)	0.02	(0.004: 0.043)	0.04	(-0.014: 0.100)
Shuttle run (s)	-0.84	(-1.554: -0.126)	-4.00	(-6.722: -1.272)
Jump height (m)	0.20	(0.051: 0.345)	0.77	(0.015: 1.534)
Grip strength (kg)	-0.07	(-0.192: 0.050)	-0.36	(-1.465: 0.753)
Fitness latent score (z-scores)	0.04	(0.005: 0.071)	0.14	(0.044: 0.242)
Girls	β	(95% CI)	β	(95% CI)
Andersen test (m)	0.04	(0.020: 0 .059)	0.11	(0.061: 0.157)
Shuttle run (s)	-1.42	(-2.191: -0.639)	-6.08	(-8.285: -3.879)
Jump height (m)	0.46	(0.239: 0.672)	0.65	(-0.153: 1.461)
Grip strength (kg)	-0.17	(-0.453: 0.115)	1.05	(-0.208: 2.299)
Fitness latent score (z-scores)	0.06	(0.021: 0.091)	0.24	(0.160: 0.329)

Note: Sex, age, school type (intervention and control), class, mother's educational level and height controlled all the analyses.

Legend: (m) refers to meters; (s) refers to seconds; (kg) refers to kilograms; (SD) refers to standard deviations; Stand. β refers to standardized Beta coefficients between the exposures and Danish language scores.

Supplemental material 3. Parameter estimates of the slope of the total and mediated influence of the fitness tests and the fitness latent score (Fitness) on Math scores for all participants and stratified by sex.

	Math scores			
All participants	Mediated by WC		Total impact	
	β	(95% CI)	β	(95% CI)
Andersen test (m)	0.02	(< -0.001: 0.036)	0.15	(0.098: 0.195)
Shuttle run (s)	-0.79	(-1.523: -0.057)	-7.00	(-9.103: -4.890)
Jump height (m)	0.29	(0.113: 0.468)	1.75	(1.017: 2.484)
Grip strength (kg)	0.06	(-0.124: 0.234)	0.84	(-0.253: 1.932)
Fitness latent score (z-scores)	0.03	(-0.006: 0.056)	0.30	(0.224: 0.381)
Boys	β	(95% CI)	β	(95% CI)
Andersen test (m)	0.01	(-0.016: 0.036)	0.16	(0.087: 0.232)
Shuttle run (s)	-0.40	(-1.300: 0.499)	-6.68	(-9.838: -3.522)
Jump height (m)	0.18	(-0.039: 0.401)	1.50	(0.458: 2.548)
Grip strength (kg)	0.07	(-0.122: 0.270)	-0.08	(-1.561: 1.403)
Fitness latent score (z-scores)	0.01	(-0.028: 0.053)	0.28	(0.166: 0.399)
Girls	β	(95% CI)	β	(95% CI)
Andersen test (m)	0.03	(< 0.001: 0.051)	0.13	(0.069: 0.201)
Shuttle run (s)	-1.23	(-2.382: -0.080)	-7.23	(-10.064: -4.397)
Jump height (m)	0.42	(0.128: 0.706)	2.06	(1.016: 3.110)
Grip strength (kg)	-0.001	(-0.333: 0.329)	1.92	(0.268: 3.570)
Fitness latent score (z-scores)	0.04	(-0.009: 0.086)	0.32	(0.213: 0.427)

Note: Sex, age, school type (intervention and control), class, mother's educational level and height controlled all the analyses.

Legend: (m) refers to meters; (s) refers to seconds; (kg) refers to kilograms; (SD) refers to standard deviations; Stand. β refers to standardized Beta coefficients between the exposures and Math scores.