Medicine & Science IN Sports & Exercise
The Official Journal of the American College of Sports Medicine
www.acsm-msse.org

. . . Published ahead of Print

Physical Activity and Academic Performance: Genetic and Environmental Associations

Sari Aaltonen[1,2], Antti Latvala[1], Aline Jelenkovic[3,4], Richard J. Rose[5], Urho M. Kujala[6], Jaakko Kaprio[1,4], Karri Silventoinen[2]

[1]Institute for Molecular Medicine (FIMM), University of Helsinki, Helsinki, Finland; [2]Department of Social Research, University of Helsinki, Helsinki, Finland; [3]Department of Physiology, Faculty of Medicine and Nursing, University of the Basque Country, Leioa, Bizkaia, Spain; [4]Department of Public Health, University of Helsinki, Helsinki, Finland; [5]Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN; [6]Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

Accepted for Publication: 6 August 2019

Medicine & Science in Sports & Exercise, Publish Ahead of Print
DOI: 10.1249/MSS.0000000000002124

Physical Activity and Academic Performance: Genetic and Environmental Associations

Sari Aaltonen[1,2], Antti Latvala[1], Aline Jelenkovic[3,4], Richard J. Rose[5], Urho M. Kujala[6], Jaakko Kaprio[1,4], Karri Silventoinen[2]

[1]Institute for Molecular Medicine (FIMM), University of Helsinki, Helsinki, Finland; [2]Department of Social Research, University of Helsinki, Helsinki, Finland; [3]Department of Physiology, Faculty of Medicine and Nursing, University of the Basque Country, Leioa, Bizkaia, Spain; [4]Department of Public Health, University of Helsinki, Helsinki, Finland; [5]Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN; [6]Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

Corresponding author: Sari Aaltonen Ph.D., Institute for Molecular Medicine (FIMM), University of Helsinki, P.O. Box 20 (Tukholmankatu 8), FIN-00014 University of Helsinki, Helsinki, Finland. E-mail: sari.s.aaltonen@helsinki.fi

The research work has been supported by the Academy of Finland (grant 266592 to S.A., A.J. and K.S., grant 277209 to A.L. and grants 100499, 205585, 141054, 265240, 263278 and 264146 to J.K.), the Finnish Ministry of Education and Culture (to S.A., U.M.K. and K.S.), the Juho Vainio Foundation (to S.A., U.M.K. and K.S.), and the Finnish Cultural Foundation (to S.A.). Data collection of the FinnTwin12 study has also been supported by the U.S. National Institute of Alcohol Abuse and Alcoholism (grants AA-12502, AA-00145 and AA-09203 to R.J.R.). The authors declare no conflict of interest. The results of the present study do not constitute endorsement by the American College of Sports Medicine. The authors state that the results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

ABSTRACT

Introduction Physical activity and academic performance are believed to be associated. Though both traits are partially heritable, it remains unclear whether these traits also share a genetic and/or environmental background in common. We aimed to examine to what extent leisure-time physical activity and academic performance share genetic and environmental effects from early adolescence to young adulthood. **Methods** Participants were Finnish twins (2543–2693 individuals/study wave) who reported their leisure-time physical activity at ages 12, 14, 17 and 24. Academic performance was assessed with teacher-reported grade point averages at ages 12 and 14 and by self-reported educational levels at ages 17 and 24. Bivariate quantitative genetic modeling at each age and between different ages was performed to decompose the trait correlation between academic performance and physical activity into genetic and environmental components. **Results** The trait correlations between leisure-time physical activity and academic performance were positive, but modest at most (r_{trait}=0.08–0.22 in males, and 0.07–0.18 in females). The genetic correlations between leisure-time physical activity and academic performance were higher than the trait correlations (r_A=0.17–0.43 in males, and 0.15–0.25 in females). Common genetic influences explained 43–100% of the trait correlations. Environmental influences shared by co-twins between leisure-time physical activity and academic performance were also correlated (r_C=0.27–0.54 in males, and 0.21–0.69 in females) explaining 41–100% of the trait correlations. Unique environmental influences were correlated only in females ($r_{E=}$0.10–0.15). **Conclusion** Both common genetic background and shared family environment (i.e., familial background) partially account for the associations observed between

leisure-time physical activity and academic performance. However, the estimates vary in magnitude by age.

Keywords: academic achievement; educational attainment; exercise; longitudinal; quantitative genetics; twins

INTRODUCTION

Regular physical activity (PA) is associated with several health benefits (1), but an increasing interest has also arisen in investigating the association between PA and other factors not directly related to health. One of these factors has been academic performance (AP) and several systematic reviews and meta-analyses have been conducted to examine how AP is associated with PA (2-7). Thus far, the evidence has been conflicting – previous studies have suggested a positive (3, 5, 6), negative (2, 8) or no association (3, 4, 8, 9) between these traits. Moreover, it is still under debate whether the possible association between the traits is causal (10). Lees et al. (2013) (10) provided evidence in their systematic review of randomized controlled trials that aerobic PA had a positive causal role in academic achievement. However, no statistically significant improvement had been detected in any of the individual studies included in the review, which makes the interpretation questionable. Our recent longitudinal twin study from early adolescence to young adulthood was also informative of potential causal effects between these traits (11): in contrast to Lees et al. (2013) (10), we found evidence for the fact that better AP potentially increases leisure-time physical activity (LTPA), rather than the reverse direction of association.

Numerous studies have supported a significant role of genetic factors in the regulation of both PA (12, 13) and AP (14-17). Based on the recent review by the American College of Sports Medicine sponsored roundtable committee, the heritability estimates for PA have ranged from moderate to very high (13). Finding specific genes for PA has been challenging in molecular genetic studies (12, 13, 18), but recently very large studies have started to make progress (19-21).

In terms of AP, a large meta-analysis of more than 12,000 primary school-aged twins from 6 countries (17) along with a study of 16-year-old British adolescents at the end of their compulsory education (15) have revealed the importance of genetic influences on the different aspects of the AP. Heritability estimates ranged from 44% (spelling) to 73% (reading) in these studies. This meta-analysis also revealed that the importance of genetic effects on AP differed between countries (17), which was also found by another international meta-analysis exploring the variation in the heritability of educational attainment (22). Furthermore, large-scale genome-wide association studies have identified hundreds of single nucleotide polymorphisms (SNPs) associated with educational attainment (23).

Because the existing results have been inconclusive in terms of the direction of association and limited with regard to their potential for clarifying the causality of the association between PA and AP, it is poorly known what lies behind the apparent association of these traits. It is possible that the association is causal, but there is also another possible explanation for the correlation between these traits. Very often familial factors (i.e., genetic and shared environmental factors) do not affect only one trait at a time, but rather are shared with several traits, and, thus, can create an association between the traits even in the absence of a causal effect (24). A common genetic and/or shared environmental background could also explain the inconsistent results on the direction and possible causality of the association between PA and AP.

To the best of our knowledge, no studies exist on the topic of whether PA behavior and AP share a genetic and/or environmental background. However, a few previous studies have shown that AP shares genetic factors with several aspects of health that are highly related to the level of PA,

such as waist circumference (25), self-reported health (26) and body mass index (27, 28). In these studies, the common genetic factors explained from 15% (waist circumference) to 30% (body mass index) of the associations between the traits. In a large genome-wide association study, genetic correlations between AP and many other health-related traits such as body mass index and risks of Alzheimer's disease, neuroticism and schizophrenia were observed (14). Moreover, shared genetic influences on physical function and cognition were found by a recent Chinese study (29). All this evidence is suggestive of a common genetic and/or environmental background between PA and AP as well.

A twin study design can provide fundamental information about the extent to which there may be genetic and environmental influences common to PA and AP. This may help to understand the ways in which the association between these two traits emerges. The aim of this study was to examine to what extent LTPA and AP share a genetic and environmental background from early adolescence through adolescence to young adulthood. Our focus was on LTPA, which may better reflect voluntary behavior and inherent abilities of the individual than many other aspects of PA behavior (e.g., school- or work-based PA). LTPA is also the most effective way to increase PA levels. Given the previous evidence, we hypothesize that there may be common genetic and/or environmental factors that may be partly responsible for the association between LTPA and AP.

METHODS

PARTICIPANTS

The participants of this study were drawn from the FinnTwin12 study, which is a longitudinal study of health and behavior in Finnish twins and their families (30). The twins were from the birth cohorts 1983–1987 and identified from the Central Population Registry of Finland. Thus far, four waves of the FinnTwin12 study have been completed. In the first phase, the twins and their parents completed study questionnaires when the twins were 11–12 years old. Subsequently, the twins were surveyed again at the mean ages of 14.0 (standard deviation (SD) 0.08 years, range 13.7–14.9), 17.6 (SD 0.26 years, range 17.2–19.5) and 24.2 years (SD 1.64 years, range 20.5–27.5). The data for the study have been collected through mailed questionnaires. The response rates have been high in each wave of data collection, ranging from 73% to 90%. The teachers of twins assessed the twins' behavior and AP at school at ages 12 and 14.

The total number of twins enrolled to the FinnTwin12 study was 5,600. In this study, we had data available on LTPA and AP from 4,179 twins (50% females) at age 12 years, 2,853 twins (51% females) at age 14 years, 4,190 twins (52% females) at age 17 years, and 3,156 twins (56% females) when they were 24 years old. Out of these individuals, 1,915 were full twin pairs at age 12, 1,173 at age 14, 1,925 at age 17, and 1,261 at age 24. The zygosity of these twins has been determined using a well-validated questionnaire (31).

ASSESSMENT OF ACADEMIC PERFORMANCE

School performance was assessed by teachers at school with a grade point average at the ages of 12 and 14 years. Teachers answered the following question: "What was the twin's grade point average at the end of the academic year?". The responses were categorized as follows: 1) better than 9, 2) from better than 8 to 9, 3) from better than 7 to 8, 4) from better than 6 to 7, 5) six and under, and 6) numeral grades not given. In the Finnish school system, grade point averages range from the lowest value of 4 to the highest value of 10. For the twins for whom numeral grades were not given at ages 12 (n=1,275) or 14 (n=12) years, their most likely grade point average category membership was imputed by using several other school performance variables reported by the teachers, such as spelling, reading aloud, and mathematics. These estimations did not bias the associations with LTPA (Supplemental Digital Content 1, Imputation of grade point average values, http://links.lww.com/MSS/B720).

At age 17 years, the twins themselves reported their current student status, which was used as a measure of AP at that stage of their lives. The responses for this item in our study were classified into three groups: 1) not studying currently, 2) in vocational school, and 3) in academically oriented upper secondary school. The self-reported highest educational degree achieved at the average age of 24 years was used as the measure of AP in young adulthood. For those twins who had not completed their education, we treated their ongoing studies as the final level of education. Educational attainment was classified into four groups: 1) compulsory education only, 2) vocational secondary education, 3) upper secondary education, and 4) tertiary education (university or polytechnic college). At age 17, the majority of the twins (62%) studied at upper

secondary school, and at age 24, the most common educational level achieved was tertiary education (53%).

ASSESSMENT OF LEISURE-TIME PHYSICAL ACTIVITY

Self-reported LTPA behavior was assessed for each study wave. The assessment of the twins' PA level was based on a structured question on the frequency of LTPA excluding physical education classes at school. The item was asked in exactly the same form in all study waves: "How often do you exercise or take part in sports in your leisure time?". However, the response options were somewhat different over time. At the study wave of age 12 years, there were 5 response options: 1) just about every day, 2) two to three times a week, 3) two to three times a month, 4) two to three times in six months, and 5) not at all. At ages 14 and 17, the 7 response options were: 1) just about every day, 2) four to five times a week, 3) two to three times a week, 4) about once a week, 5) one to two times a month, 6) less than once a month, and 7) not at all. In young adulthood, the options were the same as at ages 14 and 17, but one extra response option was included: "several times every day". The most often reported frequency of LTPA was two to three times a week at each age of the twins. Previous studies have shown strong correlations between the self-reported questions on the frequency, duration and intensity of LTPA sessions used in the Finnish twin cohort studies and PA data obtained by interview ($r=0.56$, $p<0.001$) (33) or by a detailed assessment of 12-month LTPA history ($r=0.73$, $p<0.001$) (34).

STATISTICAL METHODS

We began the statistical analyses by producing the means and SDs in monozygotic (MZ) and dizygotic (DZ) twins of the LTPA and AP variables using Stata 14.1 software (StataCorp, College Station, Texas, USA). Quantitative genetic modeling of the data was performed using structural equation modeling in OpenMx software (version 2.0.1), which is a package for extended structural equation modeling on the R statistical platform (35).

Because the quantitative genetic modeling is based on the different genetic relatedness of MZ and DZ co-twins (i.e., MZ co-twins have virtually the same DNA sequence, whereas DZ co-twins share, on average, 50% of their segregating genes), we started with information on the similarities of MZ and DZ twins with regard to LTPA and AP. We estimated intra-class correlation coefficients to quantify the degrees to which MZ and DZ twins resemble each other. The within-twin-pair correlations for LTPA and AP were systematically higher among MZ pairs than DZ pairs (see Table, Supplemental Digital Content 2, Intra-class correlation coefficients of LTPA and AP, http://links.lww.com/MSS/B721). This indicates that genetic factors are of importance in both traits. Most DZ twin correlations were more than half the MZ correlations, which suggests that environmental factors shared by co-twins are also important for both LTPA and AP. Further, DZ twin correlations for opposite-sex twin pairs were lower than for same-sex twin pairs suggesting that different genetic factors operate in males and females with regard to LTPA and AP.

Based on these principles, we decomposed the trait variation in LTPA and AP into three components: additive genetic variation (A), shared environmental variation (C), and unique

environmental variation (E) (36). Measurement error is included in the unique environmental variation. The proportion of variation accounted for by genetic influences (A effects) is called heritability (genetic influences correlate 1.0 in MZ and 0.5 in DZ twins). Shared environmental influences (C effects) refer to all environmental influences that make members of a twin pair alike (correlate 1.0 in both MZ and DZ twins) – usually interpreted as factors related to parents, siblings, household and neighborhood. Unique environmental influences (E effects) denote all environmental influences that make members of a twin pair unlike (uncorrelated in both MZ and DZ twins). High heritability estimates indicate a minor role for environmental influences, whereas low heritability estimates are suggestive of a greater role for environmental influences, shared or unshared, on the differences between individuals.

Genetic modeling began by comparing different univariate models to select the best-fitting model at each age to be used in the further analyses (see Table, Supplemental Digital Content 3, Univariate model-fitting statistics for the LTPA and AP variables, http://links.lww.com/MSS/B722). First, we determined whether shared environmental factors were present to explain the variation in the twins' LTPA and AP by comparing the full ACE model to the AE model at each age separately. Univariate model-fitting results revealed that there were significant differences in model fit between the full ACE and AE models in both LTPA and AP (p-values range from <0.001 to 0.004), except LTPA at ages 17 and 24 (see Table, Supplemental Digital Content 3, Univariate model-fitting statistics for the LTPA and AP variables, http://links.lww.com/MSS/B722), indicating that the simpler AE models did not describe the data as adequately as the more complex ACE model; thus, the shared environmental components were included in the bivariate models.

Subsequently, we tested whether there was evidence for sex-specific genetic factors related to LTPA and/or AP by analyzing whether the genetic correlations for opposite-sex twins could be constrained to 0.5 (i.e., the same as for same-sex DZ twins). The analysis of whether separate parameters should be estimated for males and females revealed evidence for a sex-specific genetic effect in LTPA at age 12 (p-value <0.001) and AP at ages 12, 14 and 17 (p-values from <0.001 to 0.02), indicating that there are different genetic influences on these traits in males and females at these ages (see Table, Supplemental Digital Content 3, Univariate model-fitting statistics for the LTPA and AP variables, http://links.lww.com/MSS/B722). Therefore, the ACE - models were conducted with separate parameter estimates for males and females.

Then, we tested whether there were differences in the absolute genetic and environmental variances in LTPA and AP between males and females. In addition to testing absolute variances, we further tested whether there were differences in the relative genetic and environmental variances (i.e., in standardized variances) between males and females. The results of absolute and relative genetic and environmental variances in LTPA and AP also showed some significant differences in both traits between males and females (see Table, Supplemental Digital Content 3, Univariate model-fitting statistics for the LTPA and AP variables, http://links.lww.com/MSS/B722). With regard to absolute variances, there were statistically significant differences in LTPA between the sexes at ages 12 (p=0.0008) and 17 (p=0.01) as well as in AP at ages 12 (p=0.02), 14 (p<0.001) and 24 (p=0.002) (see Table, Supplemental Digital Content 3, Univariate model-fitting statistics for the LTPA and AP variables, http://links.lww.com/MSS/B722), whereas the relative genetic and environmental variances in

LTPA were significantly different between sexes at ages 12 (p=0.009), 17 (p=0.004) and 24 (p=0.04), and in AP at age 17 (p=0.02).

Finally, bivariate Cholesky decompositions were conducted to estimate trait correlations (phenotypic correlation) as well as genetic and environmental correlations, and proportions of the trait correlations explained by genetic and environmental factors between LTPA and AP (37). The genetic and environmental correlations were estimated to see to what extent the covariation of the traits is potentially explained by the same genetic and environmental factors. The correlations were initially performed based on the univariate model-fitting results (ACE models) (see Table, Supplemental Digital Content 3, Univariate model-fitting statistics for the LTPA and AP variables, http://links.lww.com/MSS/B722). However, shared environmental correlations, for example, could not be reliably estimated at all ages due to the lack of significant shared environmental paths to LTPA. Because of this reason, we decided that in addition to full model estimates of the decomposed source of trait correlations, we would also conduct model estimates at each age that include only those genetic and environmental correlations that do not significantly reduce the fit of the initial full model.

We began by conducting the Cholesky decompositions between LTPA and AP separately at each age as well as across all ages. However, the results between non-consecutive ages (e.g., LTPA at age 12 and AP at age 24) showed insufficient statistical power to assess the trait correlations and to decompose the different sources of the trait correlations (see Table, Supplemental Digital Content 4, Trait correlations as well as the correlations between additive genetic, shared environmental and unique environmental factors for LTPA and AP,

http://links.lww.com/MSS/B723). Therefore, we chose to focus on cross-sectional results and results between consecutive ages.

ETHICS OF THE STUDY

The ethics committee of the Department of Public Health of the University of Helsinki (Helsinki, Finland), the ethics committee of the Helsinki University Central Hospital District (Helsinki, Finland) and the Institutional Review Board of Indiana University (Bloomington, Indiana, USA) approved the FinnTwin12 study protocol. All study methods were carried out in accordance with the approved guidelines. The parents of the participating twins or twins themselves as young adults provided written informed consent for participation in the study.

RESULTS

The means and SDs of LTPA and AP stratified by zygosity and sex are provided in Table 1. There were no significant differences by zygosity groups, except LTPA at age 17 in females ($p=0.003$). At each age, males reported higher frequencies for LTPA than females ($p=0.001$ to $p<0.001$), whereas females reported systematically higher levels of AP than males ($p<0.001$ at each age). Descriptive statistics are presented in Supplemental table 5 (see Table, Supplemental Digital Content 5, Descriptive statistics, http://links.lww.com/MSS/B724). The most common grade point average at ages 12 and 14 was 8–9 (47%) with the maximum grade point average being 10. The majority of the twins (62%) were studying at upper secondary school at age 17, and the most common educational attainment in young adulthood was tertiary education (53%). The most often reported frequency of LTPA was 2–3 times a week in each study wave. Among

twin's mothers and fathers, the frequency of LTPA was about the same: 6–10 times a month was the most often reported frequency for both (26% mothers, 25% fathers).

The heritability estimates of LTPA ranged from 30% (age 12) to 56% (age 24) in males and from 17% (age 12) to 43% (age 17) in females (Figure 1). For AP, the heritability estimates were similar or somewhat higher than for LTPA, ranging from 31% (ages 14 and 17) to 55% (age 12) in males and from 41% (age 14) to 66% (age 12) in females. The contribution of shared environmental influences to LTPA decreased dramatically from adolescence to young adulthood in both sexes (from 35% to 0% in males and from 53% to 5% in females), but were more stable in AP during the follow-up (from 27% to 29% in males and from 13% to 31% in females). The unique environmental influences were more important for LTPA (from 30% to 57%) than for AP (from 18% to 30%) at each age in both sexes.

Results of the decomposed source of the trait correlations within study waves and between consecutive study waves of LTPA and AP are presented in Tables 2 (males) and 3 (females). Statistically significant genetic and shared environmental correlations are summarized in Figure 2. Trait correlations for both LTPA and AP were positive and statistically significant but weak, ranging from 0.08 to 0.22 in males (Table 2) and from 0.07 to 0.18 in females (Table 3). The highest trait correlations between LTPA and AP were seen in late adolescence and young adulthood in both sexes: between LTPA at age 17 and AP at age 17 (r_{trait}=0.22) in males, and between LTPA at age 17 and AP at age 24 (r_{trait}=0.18) in females. The genetic and environmental correlations based on the full model between LTPA and AP fluctuated quite a bit over the transition period from early adolescence to young adulthood in both sexes.

Not all correlations could be reliably estimated in the initial full models that were based on the univariate model-fitting results. Therefore, without significant loss to the full models, the best-fitting models were identified. The importance of familial factors explaining the associations between LTPA and AP was highlighted by the fact that genetic and shared environmental correlations could not be dropped from the best-fitting models simultaneously without degrading the fit of the models – this was the case at each age in both sexes (Tables 2 and 3). Particularly, shared environmental correlations between LTPA and AP were shown to have a significant role over the follow-up from early adolescence to young adulthood: the range of significant shared environmental correlations between LTPA and AP was from 0.27 to 0.54 in males and from 0.21 to 0.69 in females (Figure 2, Tables 2 and 3). In these best-fitting models, the shared environmental influences explained almost half of the trait correlations observed between LTPA and AP (i.e., from 41% to 100%). The highest proportions of shared environmental influences explaining the trait correlation between LTPA and AP were found in early adolescence and adolescence in males, while in females the highest proportions were found in adolescence and young adulthood.

In males, the statistically significant best-fitting model correlations between the genetic influences for LTPA and AP were found when LTPA in early adolescence and young adulthood were involved (r_A from 0.17 to 0.43) (Figure 2 and Table 2). In females, the statistically significant genetic correlations were seen between LTPA at age 14 and AP at age 12 (r_A=0.15) and between LTPA at age 24 and AP at age 24 (r_A=0.25) (Figure 2 and Table 3). Similarly to the shared environmental influences, the genetic influences explained a substantial portion of the association between LTPA and AP – from 43% to 100% of the trait correlations between LTPA

and AP. The importance of common genetic influences explaining the trait correlation was highlighted particularly in young adulthood in both sexes, but also in early adolescence in males.

In many of the best-fitting models, unique environmental correlations could be dropped as a source of the trait correlations. Those unique environmental correlations that were still shown to have a role in the best-fitting models were estimated to be very low. Unique environmental influences explained less than half of the trait correlations between LTPA and AP (i.e., from 14% to 49%) – the highest proportions were found in females in early adolescence.

DISCUSSION

In the present study, we investigated to what extent LTPA and AP share a common genetic and environmental background from early adolescence to young adulthood using a genetically informative twin design. The results indicated that the associations between LTPA and AP were positive, but weak from early adolescence to young adulthood. The most important finding to emerge from this study was that the observed associations between LTPA and AP partly result from both overlapping genetic influences and overlapping familial environmental effects, which, however, vary in magnitude by age. Particularly in males, genetic influences explained a large part of the associations between LTPA and AP in early adolescence and young adulthood, whereas in females, shared environmental factors mainly explained the trait correlations between LTPA and AP. The longitudinal study design indicated that the common genetic and shared environmental influences explaining the associations between the consecutive ages of LTPA and AP were also substantial with few exceptions. In the bivariate twin model, a causal association between LTPA and AP should appear as overlap between all the variance components

contributing to both variables (i.e., as significant A, C, E correlations in our analyses). This was not the case for any pair of LTPA and AP variables, thus our findings are not supportive of a causal association existing between PA and AP, but rather indicate that familial factors (i.e., genetic and non-genetic factors shared by co-twins) are responsible for much of the association between LTPA and AP.

The results of the present study reinforce the suggestion of a positive association between PA and AP that has previously been proposed in several studies (3-6, 11). We were also able to confirm previous evidence of a significant contribution of genetic factors to PA behavior as well as to AP (12, 13, 15-17). In our study, the heritability estimates range from low to high in both LTPA and AP, which are somewhat inconsistent with previous studies suggesting moderate to high heritability estimates for individual differences in PA and AP (13, 15, 17). However, when comparing the results of different studies it is important to bear in mind that the associations reported between PA and AP partly reflect the associations between PA and cognitive ability (2, 38, 39). Furthermore, both PA and AP as variables are complex to define and can be measured in many ways. These methodological differences along with small sample sizes in some studies, as well as the fact that genetic influences change over time, may result in the discrepant findings.

As far as we know, previous studies have paid no attention to the role of the common genetic and environmental factors potentially affecting the association between PA and AP. However, many human behavioral traits are associated with several genetic variants (polygenic traits) and PA and AP are such complex traits (24). Because there are also previous studies that have suggested a common genetic background between several aspects of health and AP (25-28), we believe that

our results provide a valuable insight into the assessment of whether there are a potential common genetic and/or environmental background between different health-related variables and AP.

As hypothesized, the results of this study show that LTPA and AP share a common genetic and environmental background from early adolescence to young adulthood. We found positive genetic correlations between the traits, especially in males, indicating that the increase in LTPA is explained by the same genetic factors as the increase in AP. Further, the frequently seen moderate and high positive estimates of shared common environmental correlations in males and females indicated that a family environment that is conducive to LTPA also tends to increase AP. The current study found that the common genetic and environmental factors explained from 41% to 100% of the associations between LTPA and AP. These are markedly higher estimates than the previously reported proportions explaining the association between AP and different aspects of health (25-28). The differing results are likely explained by the variety of health and PA variables used in the studies. In addition, in our results, some of the confidence intervals are wide, which may indicate that our data is underpowered to decompose genetic and shared environmental influences explaining the proportions of the trait correlations between LTPA and AP.

Our study was limited by the absence of objective measures of LTPA. Even though the validity of PA questionnaires used in Finnish twins have been demonstrated (33, 34), it is likely that non-objective PA measures cannot be as accurate as objective PA measures (40). We were also able to utilize only the frequency measures of LTPA, which is not the most optimal way to measure

PA behavior, but was the only available PA measure over time in the longitudinal data set used in this study. We are aware of the fact that the lack of comprehensiveness in the assessment of LTPA dimensions may restrict the picture of the total LTPA behavior. However, it should be noted that the assessment method is unlikely to have an effect on the association between LTPA and AP, because the non-objective measure of LTPA still represents twins' PA behavior, yet not as comprehensively and objectively as it could.

Student status at 17 years of age and AP as young adults were also self-reported by the twins, thus, the possibility of errors cannot be avoided with regard to these variables either. In contrast to these subjective measurements, grade point averages in adolescence were reported by teachers when the twins were aged 12 and 14 years. The benefit of the measurement based on the evaluation of professionals is that teachers have evaluated the twins' school performance as objectively as possible. Because in Finland schools follow a national curriculum, the school systems and evaluations are based on similar principles all over the country. Moreover, the proficiency level of the teachers is rather similar, since nearly all teachers have undergone Master's level training. Despite all this, we are aware of the fact that grade point averages are not the most optimal measure of AP because they are not standardized and, thus, difficult to compare.

The representativeness of the study participants may also be a potential issue in our study due to the reason that we included only those twins who had complete data on LTPA and AP in the analyses. However, no evidence was found that those twins whose AP data was unavailable differed from twins whose AP data was available with regard to LTPA behavior except at age 12

(p=0.02). In contrast, those twins who did not report their LTPA behavior had significantly lower AP than those who reported their LTPA behavior (p-values were <0.001, 0.03 and 0.03, respectively) except at the latest study wave when twins were young adults. This might dilute the association between LTPA and AP.

A major strength of our study is that it provided an important opportunity to investigate the phenomenon cross-sectionally as well as longitudinally from early adolescence to young adulthood. Thus, we were able to quantify changes in the magnitude of genetic and environmental influences in different stages of life. This is important because genetic and environmental influences are not stable. A further strength of this study is the large sample size. The population-based dataset we used enables capturing the entire variation of LTPA behavior in the Finnish population across the lifespan from early adolescence to young adulthood. Various selection biases are also unlikely in our study because of the relatively high participation rates in the study waves and the inclusion of multiple domains in the questionnaires. Even though the generalizability of our study findings is good due to the population-based sample with relatively equal sex representation as well as high response rates, it is important to keep in mind that our findings are limited to individuals between 12 and 24 years of age. Thus, the trait correlations as well as genetic and environmental correlations between LTPA and AP at other ages are still poorly known.

In conclusion, the results of the present study suggest that the origin of the association between LTPA and AP is likely to be partly located in the common genetic background and shared family environment. Most importantly, the result of the present study challenges the assumption of a

potential causal relationship between LTPA and AP. In light of our results, it should not be thought that the increase in PA or AP would directly cause an increase in the other variable. For example, more practical implication to increase a child's PA behavior and AP may require interventions in a child's family environment rather than interventions that are directly aimed at promoting PA or AP. A favorable environment may also support a child's potential genetic predisposition to higher levels of PA and AP. For future studies, more information on the genetic- and environmental-based associations between PA and AP utilizing twin and family data are needed. The focus should not be only on adolescents and young adults but also on younger children and older adults. The success of this study in presenting evidence of shared common genetic factors between LTPA and AP should also lay the groundwork for potential molecular genetic studies in the field.

ACKNOWLEDGEMENTS

The research work has been supported by the Academy of Finland (grant 266592 to S.A., A.J. and K.S., grant 277209 to A.L. and grants 100499, 205585, 141054, 265240, 263278 and 264146 to J.K.), the Finnish Ministry of Education and Culture (to S.A., U.M.K. and K.S.), the Juho Vainio Foundation (to S.A., U.M.K. and K.S.), and the Finnish Cultural Foundation (to S.A.). Data collection of the FinnTwin12 study has also been supported by the U.S. National Institute of Alcohol Abuse and Alcoholism (grants AA-12502, AA-00145 and AA-09203 to R.J.R.).

AUTHOR CONTRIBUTIONS

R.J.R. and J.K. conceived, designed and contributed to the data collection of the FinnTwin12 cohort. S.A., J.K., U.M.K. and K.S. conceived and designed the present study. S.A. and K.S. conducted the statistical analyses and A.L., A.J., J.K offered statistical expertise. S.A. drafted the manuscript and A.L., A.J., R.J.R., J.K., U.M.K. and K.S. critically revised and edited the manuscript. All authors read and approved the final manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest. The results of the present study do not constitute endorsement by the American College of Sports Medicine.

The authors state that the results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

REFERENCES

1. Reiner M, Niermann C, Jekauc D, Woll A. Long-term health benefits of physical activity – a systematic review of longitudinal studies. *BMC Public Health.* 2013;13:813.

2. Esteban-Cornejo I, Tejero-Gonzalez CM, Martinez-Gomez D, et al. Objectively measured physical activity has a negative but weak association with academic performance in children and adolescents. *Acta Paediatr.* 2014;103(11):e501-6.

3. Mura G, Vellante M, Nardi AE, Machado S, Carta MG. Effects of school-based physical activity interventions on cognition and academic achievement: a systematic review. *CNS Neurol Disord Drug Targets.* 2015;14(9):1194-208.

4. Rasberry CN, Lee SM, Robin L, et al. The association between school-based physical activity, including physical education, and academic performance: a systematic review of the literature. *Prev Med.* 2011;52:S10-20.

5. Alvarez-Bueno C, Pesce C, Cavero-Redondo I, Sanchez-Lopez M, Garrido-Miguel M, Martinez-Vizcaino V. Academic achievement and physical activity: a meta-analysis. *Pediatrics.* 2017;140(6):e20171498.

6. Marques A, Santos DA, Hillman CH, Sardinha LB. How does academic achievement relate to cardiorespiratory fitness, self-reported physical activity and objectively reported physical activity: a systematic review in children and adolescents aged 6-18 years. *Br J Sports Med.* 2018;52(16):1039.

7. Singh AS, Saliasi E, van den Berg V, et al. Effects of physical activity interventions on cognitive and academic performance in children and adolescents: a novel combination of a systematic review and recommendations from an expert panel. *Br J Sports Med.* 2019;53(10):640-647

8. Van Dijk ML, De Groot RH, Savelberg HH, Van Acker F, Kirschner PA. The association between objectively measured physical activity and academic achievement in Dutch adolescents: findings from the GOALS study. *J Sport Exerc Psychol.* 2014;36(5):460-73.

9. Syvaoja HJ, Tammelin TH, Ahonen T, Kankaanpaa A, Kantomaa MT. The associations of objectively measured physical activity and sedentary time with cognitive functions in school-aged children. *PLoS One.* 2014;9(7):e103559.

10. Lees C, Hopkins J. Effect of aerobic exercise on cognition, academic achievement, and psychosocial function in children: a systematic review of randomized control trials. *Prev Chronic Dis.* 2013;10:e174.

11. Aaltonen S, Latvala A, Rose RJ, Kujala UM, Kaprio J, Silventoinen K. Leisure-time physical activity and academic performance: cross-lagged associations from adolescence to young adulthood. *Sci Rep.* 2016;15;6:39215.

12. de Geus EJ, Bartels M, Kaprio J, Lightfoot JT, Thomis M. Genetics of regular exercise and sedentary behaviors. *Twin Res Hum Genet*. 2014;17(4):262-71.

13. Lightfoot JT, De Geus EJ, Booth FW, et al. Biological/genetic regulation of physical activity level: consensus from GenBioPAC. *Med Sci Sports Exerc.* 2018;50(4):863-73.

14. Okbay A, Beauchamp JP, Fontana MA, et al. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature.* 2016;533(7604):539-42.

15. Shakeshaft NG, Trzaskowski M, McMillan A, et al. Strong genetic influence on a UK nationwide test of educational achievement at the end of compulsory education at age 16. *PLoS One.* 2013;8(12):e80341.

16. Rietveld CA, Medland SE, Derringer J, et al. GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science.* 2013;340(6139):1467-71.

17. de Zeeuw EL, de Geus EJ, Boomsma DI. Meta-analysis of twin studies highlights the importance of genetic variation in primary school educational achievement. *Trends in Neuroscience and Education.* 2015;4:69-76.

18. Kim J, Oh S, Min H, Kim Y, Park T. Practical issue in genome-wide association studies for physical activity. *Ann N Y Acad Sci.* 2011;1229:38-44.

19. Doherty A, Smith-Byrne K, Ferreira T, et al. GWAS identifies 14 loci for device-measured physical activity and sleep duration. *Nat Commun.* 2018;9(1):5257.

20. Hara M, Hachiya T, Sutoh Y, et al. Genomewide association study of leisure-time exercise behavior in Japanese adults. *Med Sci Sports Exerc.* 2018;50(12):2433-41.

21. Klimentidis YC, Raichlen DA, Bea J, et al. Genome-wide association study of habitual physical activity in over 377,000 UK Biobank participants identifies multiple variants including CADM2 and APOE. *Int J Obes (Lond).* 2018;42(6):1161-76.

22. Branigan AR, Mccallum KJ, Freese J. Variation in the heritability of educational attainment: an international meta-analysis. *Social Forces*. 2013;92(1):109-40.

23. Lee JJ, Wedow R, Okbay A, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet.* 2018;50(8):1112-21.

24. Chabris CF, Lee JJ, Cesarini D, Benjamin DJ, Laibson DI. The fourth law of behavior genetics. *Curr Dir Psychol Sci.* 2015;24(4):304-12.

25. Vermeiren AP, Bosma H, Gielen M, et al. Do genetic factors contribute to the relation between education and metabolic risk factors in young adults? A twin study. *Eur J Public Health.* 2013;23(6):986-91.

26. Boardman JD, Domingue BW, Daw J. What can genes tell us about the relationship between education and health? *Soc Sci Med.* 2015;127:171-80.

27. Della Bella S, Lucchini M. Education and BMI: a genetic informed analysis. *Quality & Quantity*. 2015;49:2577-93.

28. Silventoinen K, Sarlio-Lahteenkorva S, Koskenvuo M, Lahelma E, Kaprio J. Effect of environmental and genetic factors on education-associated disparities in weight and weight gain: a study of Finnish adult twins. *Am J Clin Nutr*. 2004;80(4):815-22.

29. Xu C, Zhang D, Tian X, et al. Genetic and environmental basis in phenotype correlation between physical function and cognition in aging Chinese twins. *Twin Res Hum Genet.* 2017;20(1):60-5.

30. Kaprio J. The Finnish twin cohort study: an update. *Twin Res Hum Genet*. 2013;16(1):157-62.

31. Jelenkovic A, Ortega-Alonso A, Rose RJ, Kaprio J, Rebato E, Silventoinen K. Genetic and environmental influences on growth from late childhood to adulthood: a longitudinal study of two Finnish twin cohorts. *Am J Hum Biol.* 2011;23(6):764-73.

32. Latvala A, Rose RJ, Pulkkinen L, Dick DM, Korhonen T, Kaprio J. Drinking, smoking, and educational achievement: cross-lagged associations from adolescence to adulthood. *Drug Alcohol Depend.* 2014;137:106-13.

33. Waller K, Kaprio J, Kujala UM. Associations between long-term physical activity, waist circumference and weight gain: a 30-year longitudinal twin study. *Int J Obes (Lond).* 2008;32(2):353-61.

34. Leskinen T, Waller K, Mutikainen S, et al. Effects of 32-year leisure time physical activity discordance in twin pairs on health (TWINACTIVE study): aims, design and results for physical fitness. *Twin Res Hum Genet*. 2009;12(1):108-17.

35. Boker S, Neale M, Maes H, et al. OpenMx: an open source extended structural equation modeling framework. *Psychometrika.* 2011;76(2):306-17.

36. Rijsdijk FV, Sham PC. Analytic approaches to twin data using structural equation models. *Brief Bioinform.* 2002;3(2):119-33.

37. Posthuma D, Beem AL, de Geus EJ, et al. Theory and practice in quantitative genetics. *Twin Res.* 2003;6(5):361-76.

38. Batty GD, Deary IJ, Schoon I, Gale CR. Childhood mental ability in relation to food intake and physical activity in adulthood: the 1970 British Cohort Study. *Pediatrics.* 2007;119(1):e38-45.

39. Belsky DW, Caspi A, Israel S, Blumenthal JA, Poulton R, Moffitt TE. Cardiorespiratory fitness and cognitive function in midlife: neuroprotection or neuroselection? *Ann Neurol.* 2015;77(4):607-17.

40. Helmerhorst HJ, Brage S, Warren J, Besson H, Ekelund U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *Int J Behav Nutr Phys Act.* 2012;9:103.

Supplemental Digital Content 1.docx

Imputation of grade point average values.

Supplemental Digital Content 2.docx

Intra-class correlation coefficients of LTPA and AP.

Supplemental Digital Content 3.docx

Univariate model-fitting statistics for the LTPA and AP variables.

Supplemental Digital Content 4.docx

Trait correlations as well as the correlations between additive genetic, shared environmental and unique environmental factors for LTPA and AP.

Supplemental Digital Content 5.docx

Descriptive statistics.

TABLE AND FIGURE LEGENDS

Table 1 Means and standard deviations of leisure-time physical activity (LTPA) and academic performance (AP) in twins by zygosity and sex.

Table 2 Trait correlations (r_{trait}) as well as the correlations between additive genetic (r_A), shared environmental (r_C) and unique environmental (r_E) factors for leisure-time physical activity (LTPA) and academic performance (AP) with 95% confidence intervals (CI) for the full and best-fitting models at ages 12, 14, 17 and 24 in males. The table presents cross-sectional correlations and correlations between consecutive ages.

Table 3 Trait correlations (r_{trait}) as well as the correlations between additive genetic (r_A), shared environmental (r_C) and unique environmental (r_E) factors for leisure-time physical activity (LTPA) and academic performance (AP) with 95% confidence intervals (CI) for the full and best-fitting models at ages 12, 14, 17 and 24 in females. The table presents cross-sectional correlations and correlations between consecutive ages.

Figure 1 The relative contributions of genetic and environmental factors to variances in leisure-time physical activity (LTPA) and academic performance (AP) with 95% confidence intervals in parentheses in males and females.

Figure 2 The statistically significant correlations between additive genetic and shared environmental factors for leisure-time physical activity (LTPA) and academic performance (AP) for the best fitting models at ages 12, 14, 17 and 24 in males and females. Line weights denote the strength of correlations: dashed thin line (--) denotes a correlation between 0 and ±0.19; solid thin line (−) denotes a correlation between ±0.2 and ±0.29; dashed thick line (--) denotes a correlation between ±0.3 and ±0.39; and solid thick line (−) denotes a correlation over ±0.4.

Figure 1

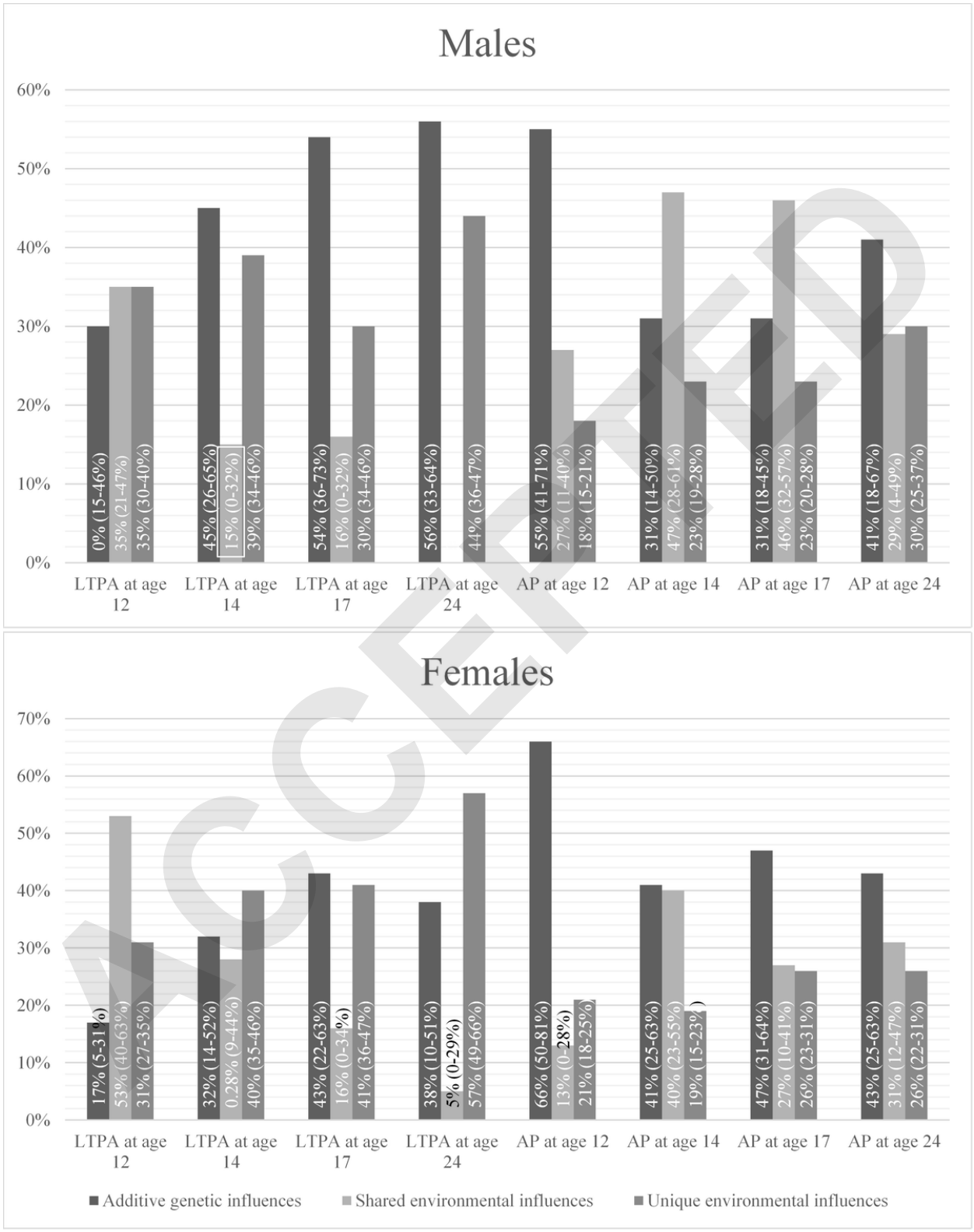

Figure 2

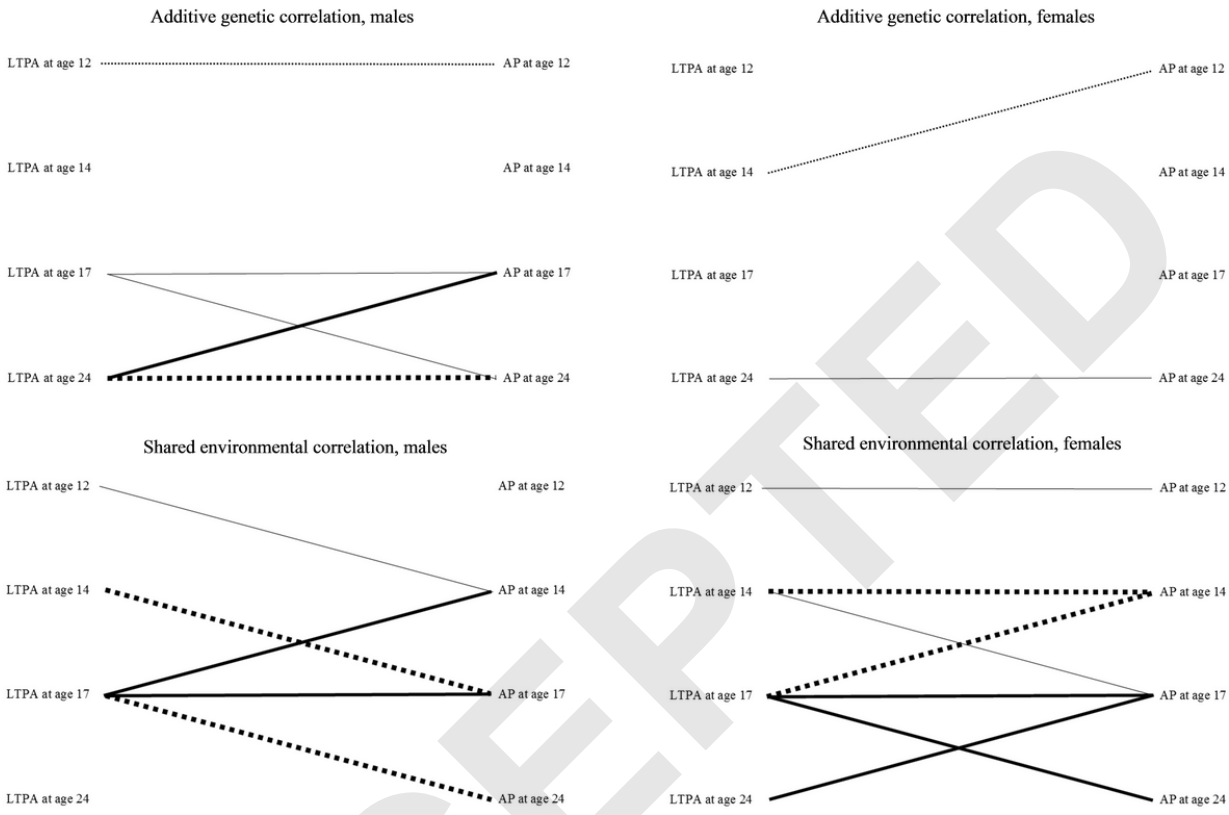

Table 1 Means and standard deviations of leisure-time physical activity (LTPA) and academic performance (AP) in twins by zygosity and sex.

Variables		MZ twins		DZ twins same sex		DZ twins opposite sex	
	Sex	Mean (SD)	Number of twin individuals	Mean (SD)	Number of twin individuals	Mean (SD)	Number of twin individuals
LTPA at age 12	♂	3.49 (1.37)	797	3.49 (1.37)	850	3.48 (1.35)	799
	♀	2.85 (1.47)	851	2.97 (1.43)	745	2.92 (1.46)	789
LTPA at age 14	♂	5.06 (1.60)	722	5.07 (1.47)	771	5.04 (1.50)	725
	♀	4.94 (1.57)	786	4.95 (1.51)	639	4.89 (1.50)	754
LTPA at age 17	♂	4.95 (1.71)	636	4.78 (1.71)	681	4.92 (1.61)	634
	♀	4.87 (1.71)	753	4.83 (1.56)	650	4.91 (1.47)	667
LTPA at age 24	♂	4.73 (1.49)	476	4.68 (1.49)	444	4.64 (1.57)	457
	♀	4.92 (1.53)	659	4.94 (1.48)	547	4.99 (1.46)	578
AP at age 12	♂	3.34 (0.68)	692	3.30 (0.69)	712	3.32 (0.73)	662
	♀	3.60 (0.64)	728	3.62 (0.68)	642	3.63 (0.66)	673
AP at age 14	♂	3.28 (0.90)	466	3.24 (0.85)	466	3.27 (0.88)	439
	♀	3.81 (0.76)	506	3.69 (0.79)	446	3.69 (0.80)	481
AP at age 17	♂	1.51 (0.54)	634	1.50 (0.58)	673	1.46 (0.58)	636

	♀	1.67 (0.56)	752	1.63 (0.60)	653	1.63 (0.57)	667
AP at age 24	♂	2.02 (1.00)	466	2.02 (1.01)	428	2.06 (1.01)	449
	♀	2.27 (0.92)	629	2.27 (0.92)	503	2.27 (0.95)	552

MZ=monozygotic, DZ=dizygotic, SD=standard deviation

Table 2 Trait correlations (r_{trait}) as well as the correlations between additive genetic (r_A), shared environmental (r_C) and unique environmental (r_E) factors for leisure-time physical activity (LTPA) and academic performance (AP) with 95% confidence intervals (CI) for the full and best-fitting models at ages 12, 14, 17 and 24 in males. The table presents cross-sectional correlations and correlations between consecutive ages.

					Trait correlation	Additive genetic correlation		Shared environmental factors		Unique environmental factors	
Trait 1	Trait 2	Model	Δdf	p-value	r_{trait} (95% CI)	r_A (95% CI)	% Explained of r_{trait}	r_C (95% CI)	% Explained of r_{trait}	r_E (95% CI)	% Explained of r_{trait}
Males											
LTPA age 12	AP age 12	Full model	-	-	0.08 (0.03 to 0.13)	0.22 (-0.20 to 0.52)	*	-0.10 (-0.61 to 0.78)	*	0.01 (-0.10 to 0.13)	*
		Best fitting model	2	0.92	0.09 (0.04 to 0.13)	0.17 (0.08 to 0.28)	100%	-	-	-	-
LTPA age 12	AP age 14	Full model	-	-	0.08 (0.02 to 0.12)	-0.07 (-0.38 to 0.21)	*	0.33 (0.03 to 0.62)	*	0.01 (-0.11 to 0.13)	*
		Best fitting model	1	0.86	0.08 (0.03 to 0.14)	-	-	0.27 (0.09 to 0.46)	100%	-	-
LTPA age 14	AP age 12	Full model	-	-	0.12 (0.08 to 0.17)	0.16 (-0.09 to 0.44)	64%	0.22 (-1.00 to 1.00)	35%	0.005 (-0.10 to 0.11)	1%
		Best fitting model	1	0.93	0.12 (0.08 to 0.17)	0.16 (-0.06 to 0.41)	66%	0.21 (-1.00 to 1.00)	34%	-	-
LTPA age 14	AP age 14	Full model	-	-	0.13 (0.07 to 0.18)	0.23 (-0.21 to 0.80)	*	0.09 (0.00 to 1.00)	*	-0.009 (-0.15 to 0.13)	*
		Best fitting model	1	0.97	0.13 (0.07 to 0.18)	0.15 (-0.09 to 0.42)	47%	0.25 (-0.12 to 1.00)	53%	-	-
LTPA age 14	AP age 17	Full model	-	-	0.10 (0.05 to 0.15)	0.10 (-0.20 to 0.36)	*	0.27 (-0.05 to 1.00)	*	-0.03 (-0.14 to 0.08)	*
		Best fitting model	2	0.72	0.10 (0.05 to 0.14)	-	-	0.36 (0.17 to 1.00)	100%	-	-
LTPA age 17	AP age 14	Full model	-	-	0.20 (0.14 to 0.26)	0.09 (-0.20 to 0.35)	18%	0.56 (0.22 to 1.00)	76%	0.04 (-0.10 to 0.19)	6%
		Best fitting model	1	0.56	0.20 (0.14 to 0.26)	0.13 (-0.10 to 0.35)	27%	0.54 (0.21 to 1.00)	73%	-	-
LTPA age 17	AP age 17	Full model	-	-	0.22 (0.17 to 0.26)	0.19 (-0.04 to 0.39)	37%	0.50 (0.20 to 1.00)	59%	0.03 (-0.08 to 0.14)	4%

		Best fitting model	1	0.58	0.22 (0.17 to 0.26)	0.22 (0.03 to 0.39)	43%	0.48 (0.19 to 1.00)	57%	-	-
LTPA age 17	AP age 24	Full model	-	-	0.21 (0.15 to 0.27)	0.22 (-0.04 to 0.48)	45%	0.40 (0.06 to 1.00)	45%	0.07 (-0.07 to 0.20)	10%
		Best fitting model	1	0.32	0.21 (0.15 to 0.27)	0.29 (0.07 to 0.51)	59%	0.36 (0.03 to 1.00)	41%	-	-
LTPA age 24	AP age 17	Full model	-	-	0.19 (0.13 to 0.25)	0.26 (-0.06 to 0.61)	*	1.00 (-1.00 to 1.00)	*	-0.03 (-0.17 to 0.12)	*
		Best fitting model	2	0.20	0.18 (0.12 to 0.24)	0.43 (0.28 to 0.63)	100%	-	-	-	-
LTPA age 24	AP age 24	Full model	-	-	0.17 (0.11 to 0.22)	0.21 (-0.09 to 0.50)	58%	1.00 (-1.00 to 1.00)	26%	0.07 (-0.06 to 0.21)	16%
		Best fitting model	2	0.45	0.16 (0.11 to 0.22)	0.34 (0.20 to 0.54)	100%	-	-	-	-

*=cannot be calculated

Table 3 Trait correlations (r_{trait}) as well as the correlations between additive genetic (r_A), shared environmental (r_C) and unique environmental (r_E) factors for leisure-time physical activity (LTPA) and academic performance (AP) with 95% confidence intervals (CI) for the full and best-fitting models at ages 12, 14, 17 and 24 in females. The table presents cross-sectional correlations and correlations between consecutive ages.

					Trait correlation	Additive genetic correlation		Shared environmental factors		Unique environmental factors	
Trait 1	Trait 2	Model	Δdf	p-value	r_{trait} (95% CI)	r_A (95% CI)	% Explained of r_{trait}	r_C (95% CI)	% Explained of r_{trait}	r_E (95% CI)	% Explained of r_{trait}
Females											
LTPA age 12	AP age 12	Full model	-	-	0.09 (0.04 to 0.14)	0.01 (-0.41 to 0.27)	5%	0.25 (0.04 to 0.85)	67%	0.10 (0.00 to 0.20)	28%
		Best fitting model	1	0.93	0.09 (0.05 to 0.14)	-	-	0.26 (0.09 to 0.56)	72%	0.10 (0.01 to 0.19)	28%
LTPA age 12	AP age 14	Full model	-	-	0.07 (0.01 to 0.13)	0.01 (-0.35 to 0.35)	6%	0.07 (-0.12 to 0.26)	47%	0.14 (0.02 to 0.26)	47%
		Best fitting model	1	0.93	0.07 (0.01 to 0.13)	-	-	0.08 (-0.04 to 0.20)	51%	0.15 (0.04 to 0.25)	49%
LTPA age 14	AP age 12	Full model	-	-	0.09 (0.04 to 0.14)	0.11 (-0.11 to 0.32)	58%	0.10 (-0.27 to 0.48)	21%	0.06 (-0.04 to 0.17)	21%
		Best fitting model	1	0.56	0.09 (0.04 to 0.14)	0.15 (0.03 to 0.31)	82%	-	-	0.05 (-0.05 to 0.16)	18%
LTPA age 14	AP age 14	Full model	-	-	0.12 (0.06 to 0.18)	-0.41 (-0.84 to 0.16)	*	0.98 (0.17 to 1.00)	*	0.12 (0.00 to 0.24)	28%
		Best fitting model	2	0.25	0.12 (0 07 to 0.17)	-	-	0.36 (0.20 to 0.66)	100%	-	-
LTPA age 14	AP age 17	Full model	-	-	0.11 (0.07 to 0.16)	0.15 (-0.04 to 0.36)	51%	0.21 (0.00 to 0.47)	49%	0.00 (-0.10 to 0.10)	0%
		Best fitting model	1	0.98	0.11 (0.07 to 0.16)	0.15 (-0.01 to 0.33)	51%	0.21 (0.00 to 0.47)	49%	-	-
LTPA age 17	AP age 14	Full model	-	-	0.11 (0.05 to 0.17)	0.07 (-0.22 to 0.31)	26%	0.28 (-0.09 to 1.00)	64%	0.04 (-0.09 to 0.16)	10%
		Best fitting model	1	0.59	0.10 (0.05 to 0.16)	-	-	0.35 (0.12 to 1.00)	86%	0.05 (-0.05 to 0.16)	14%
LTPA age 17	AP age 17	Full model	-	-	0.12 (0.07 to 0.16)	0.03 (-0.18 to 0.20)	10%	0.48 (0.20 to 1.00)	87%	0.01 (-0.08 to 0.11)	3%
		Best fitting model	2	0.88	0.11 (0.07 to 0.15)	-	-	0.51 (0.29 to 1.00)	100%	-	-

LTPA age 17	AP age 24	Full model	-	-	0.18 (0.13 to 0.24)	-0.04 (-0.37 to 0.21)	*	0.64 (0.26 to 1.00)	*	0.15 (0.04 to 0.26)	*
		Best fitting model	1	0.76	0.19 (0.13 to 0.24)	-	-	0.60 (0.32 to 1.00)	75%	0.14 (0.04 to 1.00)	25%
LTPA age 24	AP age 17	Full model	-	-	0.08 (0.03 to 0.13)	-0.04 (-0.30 to 0.18)	*	0.67 (0.10 to 1.00)	*	0.05 (-0.07 to 0.16)	*
		Best fitting model	2	0.73	0.07 (0.03 to 0.12)	-	-	0.69 (0.18 to 1.00)	100%	-	-
LTPA age 24	AP age 24	Full model	-	-	0.10 (0.05 to 0.15)	0.12 (-0.21 to 0.70)	45%	0.29 (-1.00 to 1.00)	39%	0.04 (-0.07 to 0.15)	17%
		Best fitting model	2	0.66	0.10 (0.05 to 0.15)	0.25 (0.11 to 0.53)	100%	-	-	-	-

*=cannot be calculated

Supplemental Digital Content 1

Imputation of grade point average values

A standard grading system is used in Finland, but not all schools give numerical grades during the first comprehensive school years. In the present sample, teachers of 1,249 twin participants at age 12 reported that numerical grades were not given. For these participants we imputed their most likely grade point average category based on several school performance measures reported by the teachers, as was done in earlier studies[1,2]. We used a total of seven academic and behavioural measures (i.e., spelling, writing essays, reading aloud, comprehension, mathematics, diligence, and attentiveness), reported by the teachers, as predictors of the ordinal grade point average variable in a multinomial logistic regression model. Correlations between these measures and grade point average were between 0.59 and 0.69. Based on the pseudo-R2 statistic, the measures collectively explained 44% of the variation in the grade point average variable in the multinomial logistic model.

We used the post-estimation command with the outcome option in Stata 12 (StataCorp, College Station, Texas)[3] to estimate the probability of each category and then selected the most likely category for each individual with a missing grade point average value. Of the 1,249 twin participants, 616 participants (49.3%) were imputed as having a grade point average from better than 8 to 9, 568 participants (45.5%) as having a grade point average from better than 7 to 8, and 65 participants (5.2%) as having a grade point average from better than 6 to 7. These proportions were in line with the grade point average distribution in the full sample where the corresponding proportions were 46.5%, 40.0% and 8.3%,

respectively. The predicted probabilities of belonging to these categories ranged from 0.83 to 0.93.

To check for any potential bias introduced by the imputation, we compared the correlations of the original (non-imputed) and the final (imputed and non-imputed) grade point average variables with leisure-time physical activity measures at age 12. The correlations were similar (0.03 and 0.04, respectively), suggesting that no bias was present.

References

1. Latvala, A., et al. Drinking, smoking, and educational achievement: cross-lagged associations from adolescence to adulthood. *Drug Alcohol Depend.* 137, 106–113 (2014).

2. Aaltonen S, Latvala A, Rose RJ, Kujala UM, Kaprio J, Silventoinen K. Leisure-time physical activity and academic performance: cross-lagged associations from adolescence to young adulthood. Sci Rep. 2016;15;6:39215.

3. StataCorp. Stata Statistical Software: Release 12. (StataCorp LP, 2011).

Supplemental Digital Content 2 Intra-class correlation coefficients of leisure-time physical activity (LTPA) and academic performance (AP) in twin pairs by zygosity and sex.

Variables	Sex	Intra-class correlation					
		MZ pairs		DZ same sex pairs		DZ opposite sex pairs	
		r^2 (95% CI)	Number of twin pairs	r^2 (95% CI)	Number of twin pairs	r^2 (95% CI)	Number of twin pairs
LTPA at age 12	♂	0.66 (0.60–0.71)	393	0.51 (0.43–0.57)	419	0.28 (0.22–0.35)	779
	♀	0.70 (0.65–0.75)	422	0.60 (0.53–0.66)	366		
LTPA at age 14	♂	0.63 (0.57–0.69)	354	0.37 (0.28–0.46)	377	0.20 (0.13–0.27)	710
	♀	0.62 (0.55–0.68)	388	0.43 (0.34–0.52)	339		
LTPA at age 17	♂	0.71 (0.66–0.77)	310	0.45 (0.35–0.53)	325	0.20 (0.12–0.27)	623
	♀	0.64 (0.57–0.69)	368	0.37 (0.27–0.46)	318		
LTPA at age 24	♂	0.56 (0.46–0.64)	204	0.22 (0.07–0.36)	174	0.15 (0.05–0.25)	400
	♀	0.45 (0.35–0.53)	306	0.24 (0.12–0.35)	249		
AP at age 12	♂	0.81 (0.77–0.84)	339	0.53 (0.44–0.60)	346	0.47 (0.41–0.53)	650
	♀	0.78 (0.74–0.82)	359	0.48 (0.39–0.56)	314		
AP at age 14	♂	0.78 (0.73–0.83)	210	0.61 (0.51–0.69)	201	0.47 (0.39–0.55)	375
	♀	0.80 (0.75–0.85)	232	0.61 (0.52–0.69)	194		

AP at age 17	♂	0.75 (0.70–0.79)	308	0.62 (0.55–0.68)	320	0.36 (0.29–0.42)	625
	♀	0.72 (0.67–0.77)	368	0.53 (0.44–0.60)	321		
AP at age 24	♂	0.69 (0.61–0.75)	205	0.50 (0.37–0.60)	168	0.39 (0.30–0.47)	395
	♀	0.73 (0.67–0.78)	292	0.52 (0.42–0.61)	230		

MZ=monozygotic, DZ=dizygotic, CI=confidence intervals