

Genetic Influences on Physical Activity in Young Adults: A Twin Study

LINDA MUSTELIN¹, JESSICA JOUTSI¹, ANTTI LATVALA^{1,2}, KIRSI H. PIETILÄINEN^{1,3,4,5}, AILA RISSANEN^{3,4}, and JAAKKO KAPRIO^{1,2,5}

¹Hjelt Institute, Twin Research Unit, University of Helsinki, Helsinki, FINLAND; ²Department of Mental Health and Substance Abuse Services, National Institute for Health and Welfare, Helsinki, FINLAND; ³Obesity Research Unit, Department of Medicine, Division of Internal Medicine, Helsinki University Central Hospital, Helsinki, FINLAND; ⁴Department of Psychiatry, Helsinki University Central Hospital, Helsinki, FINLAND; and ⁵Institute for Molecular Medicine, University of Helsinki, Helsinki, FINLAND

ABSTRACT

MUSTELIN, L., J. JOUTSI, A. LATVALA, K. H. PIETILÄINEN, A. RISSANEN, and J. KAPRIO. Genetic Influences on Physical Activity in Young Adults: A Twin Study. *Med. Sci. Sports Exerc.*, Vol. 44, No. 7, pp. 1293–1301, 2012. **Purpose:** The aim of this study was to investigate genetic and environmental influences on different aspects of physical activity in young adult twins. **Methods:** We studied 1274 Finnish twins with a mean age of 22.4 yr, from the population-based FinnTwin12 study. Physical activity was assessed with the Baecke Questionnaire, yielding four indexes: the sport index, leisure time activity index, work index, and total score. Quantitative genetic analyses based on linear structural equations were used to estimate the contribution of genetic and environmental factors on these physical activity traits. **Results:** The overall heritability estimates were 64% (95% confidence interval (CI) = 0.56%–0.70%) for sports activity, 41% (95% CI = 0.31%–0.51%) for leisure time activity excluding sports, 56% (95% CI = 0.48%–0.63%) for physical activity at work, and 54% (95% CI = 0.45%–0.62%) for total physical activity. Unique environmental factors accounted for the rest of the trait variances. We did not find evidence for common environmental or dominant genetic influences. The heritability estimates did not differ between men and women, and no sex-specific genetic factors were found. Sports activity and leisure time activity excluding sports were associated ($r = 0.27$), and additive genetic factors explained 57% of their association. **Conclusions:** Our results suggest that genetic factors contribute significantly to physical activity levels in young adults and that sports activity is under stronger genetic influence than leisure time physical activity excluding sports. We also concluded that physical activity at work does not seem to be associated with sports activities or other leisure time physical activity at this age. **Key Words:** GENETIC EPIDEMIOLOGY, HERITABILITY, EXERCISE BEHAVIOR, TWINS

The health benefits of physical activity are well established (47). Regular physical activity is inversely associated with many chronic diseases, such as ischemic heart disease, cerebrovascular disease, breast cancer, colorectal cancer, and diabetes mellitus (2,5,47). Physical activity is also associated with psychological well-being (32,41), and it is an important means for preventing obesity (28,32). Despite this, a large part of the population remains sedentary, making physical inactivity a health risk of epidemic proportions (14).

Genetic factors are known to account for a considerable part of the variance in physical activity within populations, but previous studies show variation in the degree of its heritability (11,15,20,25,31,41). Heritability estimates for

exercise participation ranged from 27% to 70%, with a median heritability of 62%, in a large pooled twin sample from seven countries (41). It has been suggested that genetic factors contribute more strongly to physical activity in men than in women, especially during adolescence (6,11), and that genetic influences increase from childhood to adulthood (40,42,44). A recent article shows that heritability decreases from young adulthood to age 50 (44). Thus, the heritability of physical activity would seem to be highest at the period when physical maturation is complete and physical fitness is at its best.

Twin studies are an important tool for investigating genetic and environmental influences on variation in various traits (6,8,36). Monozygotic (MZ) twins are genetically identical at the sequence level, whereas dizygotic (DZ) twins share, on average, half of their segregating genes, like ordinary siblings. Thus, a greater intrapair resemblance of MZ twins, as compared with DZ twins, indicates a contribution of genetic factors to variance in the trait studied (30).

Daily physical activities can be divided into leisure time, sports, and occupational physical activity. Heritabilities of these subdivisions are, however, poorly understood, and the heritability of occupational activity has not been assessed before. Also, previous studies have not investigated the

Address for correspondence: Linda Mustelin, Hjelt Institute, Twin Research Unit, PO Box 41, 00014 University of Helsinki, Helsinki, Finland; E-mail: linda.mustelin@helsinki.fi.

Submitted for publication March 2011.

Accepted for publication December 2011.

0195-9131/12/4407-1293/0

MEDICINE & SCIENCE IN SPORTS & EXERCISE®

Copyright © 2012 by the American College of Sports Medicine

DOI: 10.1249/MSS.0b013e3182479747

contribution of genetic and environmental factors to the associations between the different types of physical activity.

In the present study, we examined 1274 young adult MZ and DZ twins from the population-based FinnTwin12 cohort to determine the relative contribution of genetic and environmental factors on sports activities, leisure time physical activity excluding sports, and physical activity at work, as assessed by the Baecke Questionnaire (3). We also investigated whether these physical activity parameters are related to each other and whether genetic factors contribute to the associations.

MATERIALS AND METHODS

Sample. FinnTwin12 is a population-based developmental twin study of health-related behaviors and correlated risk factors (22). It consists of five consecutive birth cohorts (1983–1987) of twins identified in Finland’s Central Population Register, which covers all Finnish citizens and permanent residents (21,22). Excluded from the study were twin families in which one or both co-twins were deceased, those in which both co-twins lived apart from both biological parents, and those for which the Central Population Registry listed no residential address for a twin. Of all remaining eligible families, 87% completed the initial family questionnaire, yielding some 2800 families participating at baseline. The first survey was conducted when the twins were 11–12 yr old. A follow-up questionnaire was sent to all twins at age 14, and ~90% completed and returned it.

From this epidemiological first stage of the FinnTwin12 sample, a subset of 1035 families was selected for a second-stage intensive study. Most families were selected randomly, but the sample was also enriched with families who were considered to have a higher risk to develop alcohol problems. These families constituted 28% of the final sample. In all, 1852 twin individuals (90% of those approached) were interviewed and entered the intensive study group at age 14 and formed the target group for follow-up as young adults (Fig. 1) (21).

The follow-up of the intensive study group was conducted when the twins were 20 to 25 yr old. Altogether, 1294 twin individuals filled out the questionnaire with the Baecke scale. They were invited for a 1-d assessment in Helsinki, where they were interviewed, and filled in several questionnaires. The twins also went through clinical examinations; height, weight and waist circumference were measured; and blood samples were taken for genetic analyses. Twins who could not attend the clinical assessment filled in the questionnaires at home, were interviewed by telephone, and returned an Oragene saliva sample for DNA. The zygosity of all twins from same-sex pairs was determined by genotyping of multiple genetic markers at the Paternity Testing Unit, National Institute for Health and Welfare, using DNA from blood or saliva samples.

The study subjects provided written informed consent. The protocol was designed and performed according to the

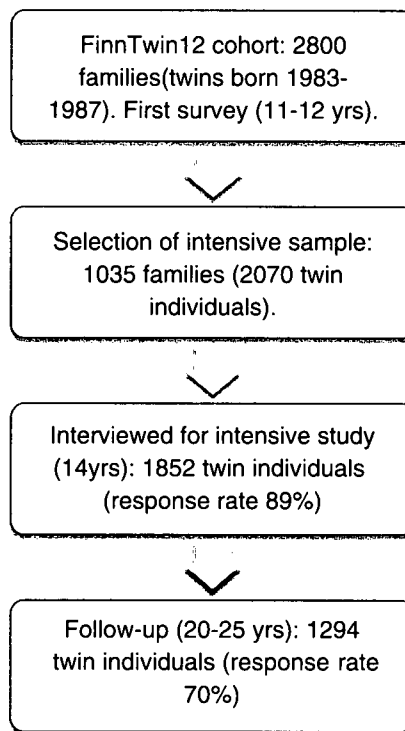


FIGURE 1—Flowchart of sample selection. The numbers refer to the actual number of families or individuals who participated in each phase. The age in years refers to the age of the twin subjects at each step of the study.

principles of the Helsinki Declaration and was approved by the Institutional Review Board of Indiana University and the Ethics Committees of the Helsinki University Department of Public Health and of the Helsinki and Uusimaa Hospital District.

The present study focuses on physical activity data from the last wave of data collection with the subjects as young adults. The final data consisted of 1274 twin individuals (686 women and 588 men), with Baecke data available. Altogether, 489 MZ individuals and 785 DZ individuals participated, including 229 (97 male and 132 female) MZ pairs and 347 (88 male, 104 female, and 155 opposite sex) DZ pairs, respectively. In addition, 31 MZ and 91 DZ twin individuals whose co-twins’ data were missing were included.

Measures. The Baecke Questionnaire (3) was used to assess physical activity. The questionnaire consists of three sections, which are sports participation, leisure time physical activity excluding sports, and work- or school-related physical activity. Each section is composed of several questions scored on a five-point scale, ranging from never to always or very often. The questions provide information on the subjects’ experience of workload and habitual physical activity. Also, one question queries the number of months per year and hours per week of participation for the two most regularly practiced sports activities, and the subjects report their main occupation. Both sports activities and occupations are scored as 1, 3, or 5 according to how physically demanding they are. Examples of “low-level” sports are sailing,

bowling, and golf; "middle-level" sports are swimming and tennis; and "high-level" sports are soccer, basketball, and boxing. Studying, office work, and driving are examples of low-level occupations; factory work, plumbing, and farming are regarded as middle-level occupations; and construction work, dock work, and professional sports are examples of high-level occupations. The low level is scored as one point, the middle level is scored as three points, and the high level is scored as five points. The leisure time activity section consists of questions querying how much time the subject spends watching television, walking, and cycling during leisure time, and the answers are scored from 1 to 5. The questionnaire has four questions on sports activity, four questions on leisure time activity excluding sports, and eight questions on work-related physical activity. The mean score of each section yields a sport index, a nonsport leisure time index, and a work index, respectively, and the sum of the three indices is called the total score. The scoring of the questionnaire items followed the standard procedure for these questions, described by Baecke et al. (3). For each of the three aforementioned indexes, the minimum score is 1, and the maximum score is 5. For the total score, the corresponding values are 3 and 15.

Statistical methods. Quantitative genetic modeling of twin data was used in the analysis (30). As mentioned earlier, MZ twins are genetically identical at the sequence level, whereas DZ twins have, on average, 50% of their genes identical by descent. Genetic variation can be divided into additive genetic (A) variation and dominance genetic variation. The former is the sum of the effects of all alleles affecting the trait, and the latter denotes interaction between alleles in the same locus. Epistatic effects, which are interaction effects between alleles in different loci, are assumed to be absent. Additive and dominance genetic effects have a correlation of 1 within MZ pairs and 0.5 and 0.25, respectively, within DZ pairs. Both MZ and DZ pairs are assumed to share the same amount of trait-relevant environmental variation, which is partly shared by members of a twin pair (common environment (C)) and partly unique to each twin individual (unique environment (E)) including any random measurement error (30). C effects result from shared factors such as dietary habits, values, leisure time activities, neighborhood, and socioeconomic situation and are expected to contribute to the correlation of both MZ and DZ twins as long as they are reared together (30), as they were in this study.

On the basis of the aforementioned assumptions, four sources of variation can be modeled: the A, the genetic dominance (D), the C, and the E components. They are estimated as standardized and latent variance components in biometrical structural equation models. To study the genetic and environmental influences on a certain trait, models based on different combinations of these components are fitted. However, when the data only include twins reared together, it is not possible to model dominance and C effects simultaneously, leading to a selection between an ACE and an ADE model (36).

The zygosity of all twins was determined from genetic markers using genomic DNA by blood or saliva samples, as mentioned, and 20 twin individuals with missing data on zygosity were excluded from the analyses. Before genetic modeling, the normality of each variable was tested, and means and SD were calculated. The Pearson correlation was used to investigate the associations between the physical activity indexes, as well as their associations with body mass index and waist circumference.

Intraclass correlations were used to investigate the relationships within twin pairs for each trait. Comparing the correlations of MZ and DZ pairs gives the first estimate of the heritability. If the intrapair correlation for MZ twins is higher than that for DZ twins, genetic effects are considered to be present.

The genetic models were carried out using the Mx statistical package using full information maximum likelihood (29). This method uses all information available in the sample including also data from twin individuals in the estimation of trait means and variances. Assumptions of twin modeling, such as equal means and variances for MZ and DZ twins and for both co-twins, were tested in saturated models, and these assumptions were met. Univariate models for sport index, leisure time activity index, work index, and total score were then fitted to estimate genetic and environmental influences and find the best-fitting model for each trait. Sex was used as a covariate in all models to remove sex-related differences in the trait means. To investigate sex differences in the genetic and environmental effects, sex limitation models for all traits were estimated, enabling the testing of both quantitative and qualitative sex differences. Quantitative sex differences refer to differences in the magnitude of A, D, C, or E influences, whereas qualitative sex differences refer to differences in the actual genetic or C factors that influence the phenotype. Qualitative genetic sex differences are likely to be present when the A correlation (r_A) among opposite-sex twin pairs is estimated to be less than 0.5 or the dominant genetic correlation (r_D) among opposite-sex twin pairs is estimated to be less than 0.25. Similarly, qualitative C sex differences are apparent when the C correlation (r_C) is estimated to be less than 1.0 among opposite-sex twin pairs. With data from twins reared together, only one of the possible qualitative sex differences can be tested at a time.

Four univariate sex limitation models were estimated consecutively for the physical activity variables. First, the full sex limitation model allowed for both quantitative and qualitative sex differences. Full models estimating either r_A or r_D (r_C) were fitted separately and compared with the Akaike information criterion (AIC), with smaller AIC values indicating a better fit. Second, the heterogeneity model allowed for quantitative sex differences and sex differences in trait variances but no qualitative sex differences, fixing the parameters r_A and r_D (r_C) to 0.5 and 0.25 (1.00), respectively. Third, the homogeneity model allowed for no sex differences in the genetic and environmental influences

TABLE 1. Means and SD in twins by sex and zygosity.

	Total	All MZ	All DZ	All Females	All Males
Number of subjects	1274	489	785	686	588
Sport index	2.98 ± 0.80	2.98 ± 0.82	2.99 ± 0.79	2.92 ± 0.78	3.06 ± 0.81**
Leisure time activity index	2.97 ± 0.63	2.94 ± 0.60	2.99 ± 0.65	3.06 ± 0.62	2.86 ± 0.63*
Work index	2.82 ± 0.67	2.82 ± 0.66	2.82 ± 0.68	2.76 ± 0.63	2.88 ± 0.72**
Total score	8.76 ± 1.30	8.74 ± 1.28	8.78 ± 1.31	8.74 ± 1.29	8.79 ± 1.30

Data are presented as mean ± SD.

* $P < 0.001$, ** $P < 0.01$, males versus females (Student's paired t -test).

or trait variances. Fourth, if the strict homogeneity model could be rejected, the scalar model was estimated allowing for sex differences in the trait variances but no qualitative or quantitative sex differences. In the scalar model, estimates for A, D (or C), and E influences are constrained to be proportionally equal for men and women. These four sex limitation models are hierarchically related (i.e., they are nested within each other in the following order: homogeneity model < scalar model < heterogeneity model < full model). After fitting the univariate sex limitation models for each of the physical activity indices, bivariate Cholesky decomposition models were estimated for each pair of correlated traits (excluding the total score). In these models, the relationship between two variables is modeled by decomposing the phenotypic covariance of the variables into proportions accounted for by A, D (or C), and E components. The degree of association of the A factors influencing the two variables can be estimated as the genetic correlation between the latent A factors for the two variables. Dominant genetic (or C) and E correlations are estimated similarly. In model comparisons, the significance of the parameters in the model was tested by dropping the parameter and evaluating the change in -2 log-likelihood between the initial model and the nested submodel using a likelihood ratio χ^2 test, where a significant ($P < 0.05$) change in χ^2 indicates that dropping the parameter significantly decreases model fit, suggesting that the parameter should be retained in the model (30).

RESULTS

Means and variances of sport index, leisure time activity index, work index, and total score are presented in Table 1. At the time of the study, 48% of the men and 36% of the women worked full-time, whereas 43% and 51%, respectively, were full-time students. However, 39% of the students reported that they also work, and 7% reported that they both study and work full-time. In addition, 5% of the women were homemakers. The remaining 9% of the men and 8% of the women either were unemployed or did not work or study

full-time for some other reason (sick leave, maternity leave, preparing for university entrance exams, military service, etc.). Mean age was 22.4 yr for both sexes (SD = 0.7 yr), and the age range was 20.9–25.1 yr for men and 20.5–26.5 yr for women.

MZ and DZ twins did not differ significantly for any of the traits. Men had a slightly higher work index (2.88 vs 2.76, for men vs women, respectively, $P < 0.01$) and sport index (3.06 vs 2.92, $P < 0.01$), whereas leisure time activity index was higher in women (2.86 vs 3.06, $P < 0.001$). Total score did not differ between the sexes. Sport index and leisure time activity index were associated ($r = 0.27$, $P < 0.001$), whereas work index showed no association with either one. Body mass index was not significantly associated with any of the physical activity parameters, whereas waist circumference was inversely and weakly associated with both sport index ($r = -0.08$, $P < 0.05$) and leisure time activity index ($r = -0.10$, $P < 0.01$) but not with work index. There were no differences in any of the physical activity measures between the randomly selected subjects and those from families with a greater risk for alcohol problems.

Intraclass correlations for sport index, leisure time activity index, work index, and total score within the twin pairs are presented in Table 2. Correlations were higher for MZ twins than for DZ twins indicating the probable effect of genetic factors. Overall, DZ correlations were less than half of the corresponding MZ correlations, suggesting dominance genetic effects. Thus, ADE models were estimated for all measures as the starting point for analyses. Correlations for opposite-sex DZ twins were often smaller than those for like-sexed DZ twins, but the 95% confidence intervals (CI) of the correlation coefficients were overlapping.

The model fit statistics and comparisons for the univariate sex limitation models are presented in Table 3. No qualitative sex differences (sex-specific genetic influences) were detected for any of the measures, and the magnitude of A, D, and E effects as well as the trait variances could be constrained equal in men and women (homogeneity model) for all traits with the exception of work index. For work index,

TABLE 2. Intraclass correlations for physical activity in twins by sex and zygosity.

	Total	MZF	MZM	DZF	DZM	OSDZ
Number of complete pairs	576	132	97	104	88	155
Sport index	0.41 (0.34–0.48)	0.65 (0.55–0.75)	0.70 (0.59–0.80)	0.26 (0.07–0.45)	0.36 (0.18–0.55)	0.11 (0.00–0.27)
Leisure time activity index	0.27 (0.19–0.34)	0.36 (0.21–0.51)	0.41 (0.24–0.58)	0.39 (0.23–0.56)	0.29 (0.10–0.49)	0.00 (0.00–0.16)
Work index	0.37 (0.30–0.45)	0.53 (0.40–0.65)	0.59 (0.46–0.72)	0.28 (0.10–0.47)	0.24 (0.04–0.44)	0.25 (0.10–0.40)
Total score	0.36 (0.28–0.43)	0.54 (0.42–0.67)	0.58 (0.44–0.71)	0.41 (0.24–0.58)	0.27 (0.07–0.47)	0.07 (0.00–0.24)

OSDZ, opposite-sex DZ.

TABLE 3. Fit statistics from univariate sex limitation models for the physical activity indexes.

Model Description	-2LL	df	AIC	Comparison	$\Delta\chi^2$	Δdf	P
Sport index							
1 Saturated model	2792.69	1219	354.69				
2 Full rA model: ADE model with quantitative and qualitative additive genetic sex differences	2803.76	1235	333.76	2 vs 1	11.06	16	0.81
3 Full rD model: ADE model with quantitative and qualitative dominant genetic sex differences	2804.45	1235	334.45	3 vs 1	11.75	16	0.76
4 Heterogeneity model: ADE model with quantitative sex differences	2804.45	1236	332.45	4 vs 2	0.69	1	0.41
5 Homogeneity model: ADE model with no sex differences	2810.40	1239	332.40	5 vs 4	5.95	3	0.11
6 Homogeneity AE model: AE model with no sex differences	2814.31	1240	334.31	6 vs 5	3.91	1	0.05
7 Homogeneity E model: E model with no sex differences	2957.62	1241	475.62	7 vs 6	143.30	1	<0.001
Leisure time activity index							
1 Saturated model	2289.40	1232	-174.60				
2 Full rA model: ADE model with quantitative and qualitative additive genetic sex differences	2314.92	1248	-181.08	2 vs 1	25.52	16	0.06
3 Full rD model: ADE model with quantitative and qualitative dominant genetic sex differences	2317.02	1248	-178.98	3 vs 1	27.62	16	0.04
4 Heterogeneity model: ADE model with quantitative sex differences	2317.02	1249	-180.98	4 vs 2	2.10	1	0.15
5 Homogeneity model: ADE model with no sex differences	2322.28	1252	-181.72	5 vs 4	5.26	3	0.15
6 Homogeneity AE model: AE model with no sex differences	2322.40	1253	-183.60	6 vs 5	0.12	1	0.73
7 Homogeneity E model: E model with no sex differences	2372.91	1254	-135.09	7 vs 6	50.50	1	<0.001
Work index							
1 Saturated model	2422.13	1226	-29.87				
2 Full rA model: ADE model with quantitative and qualitative additive genetic sex differences	2434.17	1242	-49.83	2 vs 1	12.04	16	0.74
3 Full rD model: ADE model with quantitative and qualitative dominant genetic sex differences	2434.17	1242	-49.83	3 vs 1	12.04	16	0.74
4 Heterogeneity model: ADE model with quantitative sex differences	2434.17	1243	-51.83	4 vs 2	<0.01	1	>0.99
5 Homogeneity model: ADE model with no sex differences	2444.05	1246	-47.95	5 vs 4	9.88	3	0.02
6 Scalar model: ADE model with sex differences in phenotypic variances	2435.13	1245	-54.87	6 vs 4	0.96	2	0.62
7 Scalar AE model: AE model with sex differences in phenotypic variances	2435.20	1247	-56.80	7 vs 6	0.07	1	0.79
8 Scalar E model: E model with sex differences in phenotypic variances	2540.57	1249	42.57	8 vs 7	105.37	1	<0.001
Total score							
1 Saturated model	3970.95	1193	1584.95				
2 Full rA model: ADE model with quantitative and qualitative additive genetic sex differences	3989.10	1209	1571.10	2 vs 1	18.16	16	0.31
3 Full rD model: ADE model with quantitative and qualitative dominant genetic sex differences	3990.49	1209	1572.49	3 vs 1	19.56	16	0.24
4 Heterogeneity model: ADE model with quantitative sex differences	3990.49	1210	1570.49	4 vs 2	1.39	1	0.24
5 Homogeneity model: ADE model with no sex differences	3995.51	1213	1569.51	5 vs 4	5.01	3	0.17
6 Homogeneity AE model: AE model with no sex differences	3996.97	1214	1568.97	6 vs 5	1.46	1	0.23
7 Homogeneity E model: E model with no sex differences	4091.07	1215	1661.07	7 vs 6	94.10	1	<0.001

The best-fitting models are in boldface.

LL, log-likelihood; A, additive genetic variance component; D, dominant genetic variance component; E, unique environmental variance component.

the homogeneity model could be rejected, but the scalar model had a good fit. Constraining the models further by dropping the D component did not significantly decrease model fit, but for the leisure time activity index, this test was bordering on statistical significance ($P = 0.05$). The A effects were highly significant for all traits, as expected. Thus, AE models where the A and E components accounted for equal proportions of variance in men and women fitted best for all traits, with an ADE model as an equally acceptable alternative for sport index. Estimates of the A and E variance components from the best-fitting models are given in Table 4. The heritabilities of the physical activity measures were moderately large, ranging from 41% to 64%. In the ADE model for sport index, the proportions of variance explained by A, D, and E were 25% (95% CI = 0%–64%),

41% (95% CI = 0%–71%), and 34% (95% CI = 30%–44%), respectively.

On the basis of the univariate models, we fitted a bivariate ADE Cholesky model, collapsing men and women, to estimate the relative importance of genetic and environmental factors on the relationship between sport index and leisure time activity index. All D influences could be dropped from this model ($\chi^2\Delta = 4.51$, $df = 3$, $P = 0.21$), whereas both rA ($rA = 0.32$ (95% CI = 0.18–0.45), $\chi^2\Delta = 16.58$, $df = 1$, $P < 0.001$) and E correlation ($rE = 0.27$ (95% CI = 0.15–0.38), $\chi^2\Delta = 19.10$, $df = 1$, $P < 0.001$) were highly significant. According to this model, A factors accounted for 57% (95% CI = 33%–77%) of the trait correlation between sport index and leisure time activity index, whereas E factors accounted for the rest.

TABLE 4. Additive genetic and unique environmental variance components.

	Additive Genetic Effects (A)		Unique Environmental Effects (E)	
		95% CI		95% CI
Sport index	0.64	0.56–0.70	0.36	0.30–0.44
Leisure time activity index	0.41	0.31–0.51	0.59	0.49–0.69
Work index	0.56	0.48–0.63	0.44	0.37–0.52
Total score	0.54	0.45–0.62	0.46	0.38–0.55

DISCUSSION

In the present study, we examined the genetic and environmental effects on physical activity as measured by the Baecke Questionnaire in young adult twins. The results suggest that genetic factors contribute significantly to individual differences in physical activity, with A effects explaining between 41% and 64% of the variance in the four physical activity indexes studied. Sports participation had the highest heritability, whereas the relative contribution of genes seemed to be the least for leisure time activity. E factors explained the remaining variance, whereas C and dominance genetic effects were found to be nonsignificant. Sports participation and leisure time activity excluding sports were associated, and genetic factors accounted for 57% of their relationship.

These heritability estimates on physical activity are consistent with results from previous studies, which have reported that variation due to genetic factors ranges from 27% to 71% (6,11,25,28,41). The wide range of reported heritability estimates may be explained not only by the variety of methods used to measure physical activity (6,11) but also by the various definitions of the trait (23,40). Also, sample size, as well as the sex and age of the subjects, may affect the results (40). Finally, there may be true differences in the genetic makeup of the populations and the distribution of relevant environmental conditions for the trait (45).

For assessing physical activity, we used the Baecke Questionnaire, in which physical activity is subdivided into sports participation, leisure time physical activity excluding sports, and work- or school-related physical activity. The highest heritability was found for sports participation, with genetic factors explaining 64% of the variance, whereas the lowest heritability was found for leisure time physical activity excluding sports, with genetic factors explaining 41% of its total variance. Maia et al. (25) found similar results in a study in which physical activity was likewise assessed with the Baecke Questionnaire. The study consisted of 411 Portuguese twin pairs of both sexes between the ages of 12 and 25 yr. The heritability estimates were 69% and 40% for sports activity and 63% and 32% for leisure time activity for men and women, respectively. No work index was calculated, and therefore, a total score, which describes the total daily physical activity, was not possible to obtain. Several studies focusing only on sports participation have also been made, and the results follow a similar pattern with heritability estimates ranging from 44% to 85% (6,40,41). The higher heritability estimates seen for sports participation compared with those for leisure time activity excluding sports could be explained by genetic factors apparently influencing participation in specific intense sports activities more than moderate activity (6,11). The heritability estimate for occupational physical activity was 56% in the present study, and occupational physical activity was not associated with the other physical activity parameters. Rombaldi et al. (38) also found that housework and occupational physical activity were not related to participation

in leisure time physical activity. We found no previous studies investigating the heritability of occupational physical activity.

In this study, the heritability of total physical activity was 54%. This is in line with previous studies on daily physical activity (6,11,15,20,42). In most studies, E effects contribute to the rest of the variance leaving the effect of C nonsignificant (11,15). However, in a small study ($N = 40$ twin pairs) by Joosen et al. (20), C and E factors explained all of the variance in physical activity recorded with a triaxial accelerometer in a respiration chamber for 24 h, and no genetic contribution was found. Energy expenditure was measured simultaneously and presented a similar pattern. In 2-wk daily life measurements, however, the A contribution to physical activity was 78% (95% CI = 57%–87%) with E factors explaining the rest of the variance. Duncan et al. (15) also presented results differing slightly from the majority of observations. In a sample consisting of 1003 same-sex twin pairs (62% women) with a mean age of 30 yr living in Washington State, E provided the strongest influence on physical activity, with genetic factors accounting for only 11% to 45% of the total variance. Physical activity was measured by the subjects' self-reported amount of total minutes per week of moderate-to-vigorous activity.

There are many possible routes through which genetics might influence physical activity. Many traits related to exercise capacity, such as body composition, body type, and somatic dimensions, as well as aerobic power, muscle strength, and muscle endurance, are influenced moderately or highly by genetic factors (7,9,10,27,39). Because perceived exercise competence is associated with physical activity and predicts both exercise behavior and cardiorespiratory fitness in adolescents (16,26), these physical characteristics may directly or indirectly influence physical activity. Our previous finding, that there are genetic associations between sports participation and both aerobic capacity and body composition, supports this (27). Personality characteristics, like extraversion, neuroticism, and conscientiousness, also affect activity behavior (12,37), and because of their relatively high heritability (19), they are likely to contribute to the genetic influence on physical activity. Occupational physical activity seems to be unrelated to leisure time physical activity and might reflect educational level because occupations containing heavy physical work seldom require an academic education. Educational achievement is influenced by genetics (4,46), as are many attitudes and personality characteristics (17,24,33), which may influence the choice of occupation. However, a major portion of this sample is still working on their educational achievement, and therefore, the heritability estimate of the work index might be quite different when all the subjects are in the labor force. Further, the heritability of work index may be modified by the type of work in such a situation.

Only few data are available on specific genes associated with exercise behavior and physical activity. The first genome-wide association study on physical activity level

was performed in 2009 in a Dutch-American twin sample of 2622 subjects (13). Study subjects were divided into “exercisers” and “nonexercisers” on the basis of their answers on questions about leisure time physical activity. It was found that none of the 1.6 million SNPs reached the commonly used $P = 5 \times 10^{-8}$ threshold for genome-wide significance, nor were previously reported candidate gene associations (*ACE*, *CASR*, *CYP19A1*, *DRD2*, *LEPR*, and *MC4R* genes) or linkage findings replicated. As has been the case for many other traits with complex genetics, much larger studies are obviously needed before we can draw any definitive conclusions about the specific genes having an effect on physical activity.

That C factors did not influence young adults’ physical activity was expected. Studies made among children, adolescents, and adults show that C influences on physical activity are almost completely restricted to children and young adolescents (40–42), whereas genetic factors seem to be of little importance in explaining individual differences up till age 13–14. Genetic factors start to gain importance in late adolescence, and the role of C decreases simultaneously. A peak of genetic effects is seen at the age of 19–20, after which they start to decrease, reaching a stable 50% contribution of total variance in physical activity in middle-age adults (42).

Men had higher mean scores of the sport and work indexes, whereas women had higher scores for leisure time activity excluding sports. That men received higher scores in sports participation is not surprising because they tend to participate in more physically demanding sports than women and usually also spend more time exercising (6,23,40,41). A possible explanation for the higher work index of men is that a significantly greater amount of women were studying full-time as compared with men (52% vs 44%), placing them in the lowest activity group in work index calculation. Men may, in addition, be more interested in jobs involving heavy physical work or are able to undertake heavy work as compared with women (43).

Although sex differences were present in the means of the indices, we found no evidence for different heritability patterns in men and women, unlike some previous studies suggesting higher heritability of physical activity in men (6,11). It is possible that not enough statistical power was available in this sample size to recognize sex-specific genetic effects, which is suggested by the fact that despite smaller intraclass correlations among opposite-sex twins compared with like-sexed DZ twins, homogeneity or scalar models fitted to the data best. There is, however, no previous evidence for different physical activity genes acting in men and women (6,25), although a study made by Stubbe et al. (41) suggests that the genetic factors influencing exercise participation do not entirely overlap between the sexes.

This study has several strengths. No previous study has investigated the heritability of physical activity using multiple dimensions of activity. Physical activity is often inaccurately associated only with sports participation, which is,

in fact, only a fraction of the daily physical activity for the majority of the population (6). Considering the trait in a broader sense is thus of great importance. Rather than only focusing on sports participation and activity during leisure time, this study also includes activity associated with work or school. Maia et al. (25) also used the Baecke Questionnaire to study the heritability of physical activity, but no work index was calculated, apparently because of the young age of the study subjects. Therefore, a total score, which describes the total daily physical activity, was also not obtained, which makes the present study the first to use the Baecke Questionnaire as a whole in assessing the heritability of physical activity. The Baecke Questionnaire is well validated (18,34,35) and showed, in a study investigating the validity of three commonly used physical activity questionnaires, the highest correlation with physical activity level, as measured with the doubly labeled water method (35).

Further, there are only few studies where women and men have been represented this equally. The equal sex distribution ensures that the results correspond well to the overall population. In addition, heritability of physical activity in young adults has not previously been studied, and as the age range was restricted, age and cohort effects were minimized. The zygosity of the twins, which is crucial for this study, was objectively determined from genetic markers using DNA by blood or saliva samples, unlike in the majority of genetic epidemiology research on large twin samples where zygosity is queried by questionnaire, which may have some misclassification (25).

The study also has limitations. The data are based on self-report, and it is known that people tend to overrate their behavior when it comes to activity, a behavioral pattern known as social desirability (1). In this study, overreporting of physical activity may thus be present. Further, the results of this study cannot necessarily be generalized to other populations because environmental and possibly genetic differences exist across countries (15). Studies suggest, for instance, that people in northern European countries tend to exercise more than people in the southern European countries (42). The sample in this study is fairly limited consisting of only Finns. The majority of the studies measuring physical activity are, however, made in Scandinavia or in countries such as the United Kingdom, The Netherlands, Australia, or in other Western countries (15), so these results should be comparable with each other. Finally, although the twin correlations suggested the presence of dominant genetic effects, the present sample did not provide sufficient statistical power for their detection in the variance component models.

In conclusion, we found that genetic factors contributed significantly to physical activity levels in young adults and that sports activity was under stronger genetic influence than nonsport leisure time physical activity. Genetic factors accounted for 57% of the relationship between sports activity and nonsport leisure time activity, whereas neither of these was associated with occupational physical

activity. In the future, the identification of specific physical activity genes could help understand the interplay between genes and environment in physical activity behavior.

This study was funded by the Academy of Finland Center of Excellence in Complex Disease Genetics (grants 213506 and 129680), Academy of Finland (grants 44069, 100499, 118555, 141054, and 108297), the European Community's Seventh Framework Program

(FP7/2007–2013), ENGAGE Consortium grant agreement HEALTH-F4-2007-201413, and Helsinki University Central Hospital grants. Linda Mustelin was supported by the Research Foundation of the University of Helsinki, and Kirsi Pietiläinen was supported by the Novo Nordisk and Gyllenberg foundations. Linda Mustelin and Jessica Joutsi shared first authorship.

None of the authors declare any conflict of interest.

The results of this study do not constitute endorsement by the American College of Sports Medicine.

REFERENCES

1. Adams SA, Matthews CE, Ebbeling CB, et al. The effect of social desirability and social approval on self-reports of physical activity. *Am J Epidemiol*. 2005;161(4):389–98.
2. Allender S, Foster C, Scarborough P, Rayner M. The burden of physical activity-related ill health in the UK. *J Epidemiol Community Health*. 2007;61(4):344–8.
3. Baecke JA, Burema J, Frijters JE. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *Am J Clin Nutr*. 1982;36(5):936–42.
4. Baker LA, Treloar SA, Reynolds CA, Heath AC, Martin NG. Genetics of educational attainment in Australian twins: sex differences and secular changes. *Behav Genet*. 1996;26(2):89–102.
5. Berlin JA, Colditz GA. A meta-analysis of physical activity in the prevention of coronary heart disease. *Am J Epidemiol*. 1990;132(4):612–28.
6. Beunen G, Thomis M. Genetic determinants of sports participation and daily physical activity. *Int J Obes Relat Metab Disord*. 1999;23(Suppl 3):S55–63.
7. Bogl LH, Latvala A, Kaprio J, Sovijarvi O, Rissanen A, Pietiläinen KH. An investigation into the relationship between soft tissue body composition and bone mineral density in a young adult twin sample. *J Bone Miner Res*. 2011;26(1):79–87.
8. Boomsma D, Busjahn A, Peltonen L. Classical twin studies and beyond. *Nat Rev Genet*. 2002;3(11):872–82.
9. Bouchard C, Sarzynski MA, Rice TK, et al. Genomic predictors of the maximal O₂ uptake response to standardized exercise training programs. *J Appl Physiol*. 2011;110:1160–70.
10. Bray MS, Hagberg JM, Perusse L, et al. The human gene map for performance and health-related fitness phenotypes: the 2006–2007 update. *Med Sci Sports Exerc*. 2009;41(1):35–73.
11. Carlsson S, Andersson T, Lichtenstein P, Michaelsson K, Ahlbom A. Genetic effects on physical activity: results from the Swedish Twin Registry. *Med Sci Sports Exerc*. 2006;38(8):1396–401.
12. de Bruijn GJ, de Groot R, van den Putte B, Rhodes R. Conscientiousness, extroversion, and action control: comparing moderate and vigorous physical activity. *J Sport Exerc Psychol*. 2009;31(6):724–42.
13. De Moor MH, Liu YJ, Boomsma DI, et al. Genome-wide association study of exercise behavior in Dutch and American adults. *Med Sci Sports Exerc*. 2009;41(10):1887–95.
14. Dumith SC, Hallal PC, Reis RS, Kohl HW 3rd. Worldwide prevalence of physical inactivity and its association with human development index in 76 countries. *Prev Med*. 2011;53:24–8.
15. Duncan GE, Goldberg J, Noonan C, Moudon AV, Hurvitz P, Buchwald D. Unique environmental effects on physical activity participation: a twin study. *PLoS One*. 2008;3(4):e2019.
16. Gao Z. Perceived competence and enjoyment in predicting students' physical activity and cardiorespiratory fitness. *Percept Mot Skills*. 2008;107(2):365–72.
17. Horn JM, Plomin R, Rosenman R. Heritability of personality traits in adult male twins. *Behav Genet*. 1976;6(1):17–30.
18. Jacobs DR Jr, Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Med Sci Sports Exerc*. 1993;25(1):81–91.
19. Jang KL, Livesley WJ, Vernon PA. Heritability of the big five personality dimensions and their facets: a twin study. *J Pers*. 1996;64(3):577–91.
20. Joosen AM, Gielen M, Vlietinck R, Westerterp KR. Genetic analysis of physical activity in twins. *Am J Clin Nutr*. 2005;82(6):1253–9.
21. Kaprio J. Twin studies in Finland 2006. *Twin Res Hum Genet*. 2006;9(6):772–7.
22. Kaprio J, Pulkkinen L, Rose RJ. Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families. *Twin Res*. 2002;5(5):366–71.
23. Kriska AM, Caspersen CJ. Introduction to a collection of physical activity questionnaires. *Med Sci Sports Exerc*. 1997;29(Suppl 6):S5–9.
24. Luciano M, Wainwright MA, Wright MJ, Martin NG. The heritability of conscientiousness facets and their relationship to IQ and academic achievement. *Pers Individ Differ*. 2006;40(6):1189–99.
25. Maia JA, Thomis M, Beunen G. Genetic factors in physical activity levels: a twin study. *Am J Prev Med*. 2002;23(Suppl 2):87–91.
26. Moreno-Murcia JA, Hellin P, Gonzalez-Cutre D, Martinez-Galindo C. Influence of perceived sport competence and body attractiveness on physical activity and other healthy lifestyle habits in adolescents. *Span J Psychol*. 2011;14(1):282–92.
27. Mustelin L, Latvala A, Pietiläinen KH, et al. Associations between sports participation, cardiorespiratory fitness, and adiposity in young adult twins. *J Appl Physiol*. 2011;110(3):681–6.
28. Mustelin L, Silventoinen K, Pietiläinen K, Rissanen A, Kaprio J. Physical activity reduces the influence of genetic effects on BMI and waist circumference: a study in young adult twins. *Int J Obes (Lond)*. 2009;33(1):29–36.
29. *Mx: Statistical Modeling*. Richmond (VA): Department of Psychiatry; 2003.
30. Neale MC, Cardon LR. *Methodology for Genetic Studies of Twins and Families*. Dordrecht: Kluwer Academic Publisher; 1992. 496 p.
31. Nelson MC, Gordon-Larsen P, North KE, Adair LS. Body mass index gain, fast food, and physical activity: effects of shared environments over time. *Obesity (Silver Spring)*. 2006;14(4):701–9.
32. O'Donovan G, Blazevich AJ, Boreham C, et al. The ABC of Physical Activity for Health: a consensus statement from the British Association of Sport and Exercise Sciences. *J Sports Sci*. 2010;28(6):573–91.
33. Olson JM, Vernon PA, Harris JA, Jang KL. The heritability of attitudes: a study of twins. *J Pers Soc Psychol*. 2001;80(6):845–60.
34. Philippaerts RM, Lefevre J. Reliability and validity of three physical activity questionnaires in Flemish males. *Am J Epidemiol*. 1998;147(10):982–90.
35. Philippaerts RM, Westerterp KR, Lefevre J. Doubly labelled water validation of three physical activity questionnaires. *Int J Sports Med*. 1999;20(5):284–9.
36. Posthuma D, Beem AL, de Geus EJ, et al. Theory and practice in quantitative genetics. *Twin Res*. 2003;6(5):361–76.
37. Rhodes RE, Smith NE. Personality correlates of physical activity: a review and meta-analysis. *Br J Sports Med*. 2006;40(12):958–65.

38. Rombaldi AJ, Menezes AM, Azevedo AM, Hallal PC. Leisure-time physical activity: association with activity levels in other domains. *J Phys Act Health*. 2010;7(4):460–4.
39. Silventoinen K, Magnusson PK, Tynelius P, Kaprio J, Rasmussen F. Heritability of body size and muscle strength in young adulthood: a study of one million Swedish men. *Genet Epidemiol*. 2008;32(4):341–9.
40. Stubbe JH, Boomsma DI, De Geus EJ. Sports participation during adolescence: a shift from environmental to genetic factors. *Med Sci Sports Exerc*. 2005;37(4):563–70.
41. Stubbe JH, Boomsma DI, Vink JM, et al. Genetic influences on exercise participation in 37,051 twin pairs from seven countries. *PLoS One*. 2006;1:e22.
42. Stubbe JH, de Geus EJ. Genetics of exercise behavior. In: *Handbook of Behavior Genetics*. Springer; 2009.
43. Su R, Rounds J, Armstrong PI. Men and things, women and people: a meta-analysis of sex differences in interests. *Psychol Bull*. 2009;135(6):859–84.
44. Vink JM, Boomsma DI, Medland SE, et al. Variance components models for physical activity with age as modifier: a comparative twin study in seven countries. *Twin Res Hum Genet*. 2011;14(1):25–34.
45. Visscher PM, Hill WG, Wray NR. Heritability in the genomics era—concepts and misconceptions. *Nat Rev Genet*. 2008;9(4):255–66.
46. Wainwright MA, Wright MJ, Geffen GM, Luciano M, Martin NG. The genetic basis of academic achievement on the Queensland Core Skills Test and its shared genetic variance with IQ. *Behav Genet*. 2005;35(2):133–45.
47. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. *CMAJ*. 2006;174(6):801–9.

Copyright of Medicine & Science in Sports & Exercise is the property of Lippincott Williams & Wilkins and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.