Regular exercise, subjective wellbeing, and internalizing problems in adolescence: causality or genetic pleiotropy?

Meike Bartels[1,2] *, Marleen H. M. de Moor[1,2], Niels van der Aa[1,2], Dorret I. Boomsma[1,2] and Eco J. C. de Geus[1,2]

[1] Department of Biological Psychology, VU University Amsterdam, Amsterdam, Netherlands
[2] EMGO Institute for Health and Care, VU University Medical Centre, Amsterdam, Netherlands

This study tests in a genetically informative design whether exercise behavior causally influences subjective wellbeing (SWB) and internalizing problems (INT). If exercise causally influences SWB and INT, genetic and environmental factors influencing exercise behavior will also influence SWB and INT. Furthermore, within genetically identical (MZ) twin pairs, the twin who exercises more should also show higher levels of SWB and lower levels of INT, than the co-twin who exercises less, because genetic confounding is excluded. Data on these phenotypes were available in a sample of 6317 adolescent twins and 1180 non-twin-siblings. Most participants had longitudinal data with 2-year follow-up. Exercise behavior was cross-sectionally and longitudinally associated with fewer internalizing problems and increased SWB (correlations ranged from 0.12 to 0.16). Cross-sectional and longitudinal associations were mainly accounted for by genetic factors, whereas the contribution of environmental factors was negligible. Within MZ twin pairs, the twin who exercised more did not show fewer internalizing problems and increased SWB. This was found cross-sectionally and longitudinally. We conclude that exercise behavior is associated with fewer internalizing problems and higher levels of SWB. The association largely reflects the effects of common genetic factors on these traits.

Keywords: regular exercise, subjective wellbeing, internalizing behavior, causality, genetic pleiotropy, twin-sibling

INTRODUCTION

In the adult population regular exercise has consistently been associated with higher levels of wellbeing such as fewer depressive and anxious symptoms and higher levels of subjective wellbeing (SWB). Longitudinal studies showed that exercise participation at baseline predicted fewer internalizing problems at follow-up (Camacho et al., 1991; Wise et al., 2006; de Moor et al., 2008). In general, the conclusion that has been drawn from these studies is that increased levels of psychological wellbeing found in exercisers are the reflection of causal effects of exercise.

However an alternative hypothesis explaining the prospective associations between exercise behavior and psychological wellbeing may be that underlying factors influencing exercise behavior at baseline also influence psychological wellbeing at follow-up. Genetic variation among individuals may represent such an underlying factor. Under this hypothesis, prospective associations between exercise behavior and psychological wellbeing reflect the effects of common genetic factors influencing a disposition for exercise participation as well as a disposition for psychological wellbeing, a process referred to as genetic pleiotropy (de Geus and de Moor, 2008). When genetic effects on exercise behavior at baseline are associated with genetic effects on psychological wellbeing at a later time point, this leads to a longitudinal correlation between exercise behavior and psychological wellbeing that can mistakenly be taken for a causal effect of exercise.

Evidence for the hypothesis of genetic pleiotropy in exercise behavior and psychological wellbeing was supported in population-based adult twin studies (Stubbe et al., 2007; de Moor et al., 2008). de Moor et al. (2008) demonstrated cross-sectionally and longitudinally that genetic factors influencing a person's tendency to engage in exercise also influence lower risk for depressive and anxious symptoms, whereas environmental factors affecting exercise behavior do not affect depressive and anxious symptoms. Strong support for a non-causal explanation came from the finding that in genetically identical twin pairs, the twin who exercised more did not have fewer depressive and anxious symptoms than his or her co-twin who exercised less. Moreover, individuals who increased their exercise participation over time did not show a parallel decrease in depressive and anxious symptoms.

Many studies, like the above, have defined wellbeing as the absence of depressive and anxious symptoms. Stubbe et al. (2007) added to this by using happiness and satisfaction with life as an indicator of positive wellbeing. Using both internalizing and SWB as outcome measures for psychological wellbeing has the important advantage that it is sensitive to the entire spectrum of psychological wellbeing (Proctor et al., 2009). In a co-twin control study, Stubbe et al. (2007) reported that although exercisers were more satisfied with their lives and happier than non-exercisers, the odds ratios decreased and were no longer significant in MZ and DZ twin pairs discordant for exercise participation. Taken together these results indicate that the association between exercise and SWB was most likely be explained by genetic factors influencing both traits.

Most research up to date has focused on adults. It is important to further study the association between exercise behavior and psychological wellbeing during adolescence because several studies reported that the prevalence of exercise participation declines during adolescence (van Mechelen et al., 2000; Stubbe et al., 2005; van der Aa et al., 2010a), while the prevalence of internalizing problems shows its most prominent increase (Costello et al., 2005). Furthermore, the genetic architecture of exercise behavior was found to be vastly different across the life span with a peak in heritability during late adolescence (Beunen and Thomis, 1999; Maia et al., 2002; Stubbe et al., 2005; van der Aa et al., 2010a). van der Aa et al. (2010a), for instance, showed that variation in exercise behavior during adolescence was largely accounted for by genetic factors (72–85% in boys and 38–80% in girls), whereas shared environmental factors only accounted for part of the variation in young adolescent girls (46%). Among adolescents, individual differences in internalizing and SWB have also been found to be substantially influenced by genetic factors, whereas shared environmental factors accounted for variation in internalizing only during young adolescence and were absent with regard to SWB (Bartels and Boomsma, 2009; Lamb et al., 2010; Bartels et al., 2011).

In the present study we tested whether exercise behavior causally influences internalizing problems and SWB in a large population-based sample of adolescent twins and their non-twin-siblings. Cross-sectional and longitudinal data were used, providing a strong non-experimental design to test the causal hypothesis. Applying the procedure as described by de Moor et al. (2008), four predictions generated by the causal hypothesis were tested. A graphical representation of the four models used to test causal effects of exercise behavior on internalizing problems and SWB is presented in **Figure 1**.

A first prediction is that, if exercise causally influences internalizing problems and SWB, genetic and environmental factors influencing individual differences in exercise behavior will also, through the causal chain, influence individual differences in internalizing problems and SWB. This can be tested in a bivariate genetic model by computing genetic and environmental correlations between two traits. If genetic and environmental correlations are both significant this would be in line with the causal hypothesis (see **Figure 1A**). If only the genetic correlation is significant this is more compatible with the hypothesis of common genetic factors underlying exercise behavior, internalizing problems, and SWB.

A second prediction of the causal hypothesis is that monozygotic within-twin pair differences in exercise behavior should be associated with within-twin pair differences in internalizing problems and SWB. The twin who exercises more should display fewer internalizing problems and higher levels of SWB than the co-twin who exercises less (see **Figure 1B**). This would be in line with the causal hypothesis and it excludes confounding by genetic factors (the twins are genetically identical), whereas if the association is not significant this would falsify the causal effect of exercise.

A third prediction made by the causal hypothesis is that, if exercise causes fewer internalizing problems and higher levels of SWB, there should be a significant longitudinal association between exercise at baseline and internalizing problems and SWB at follow-up. This should be independent of common genetic
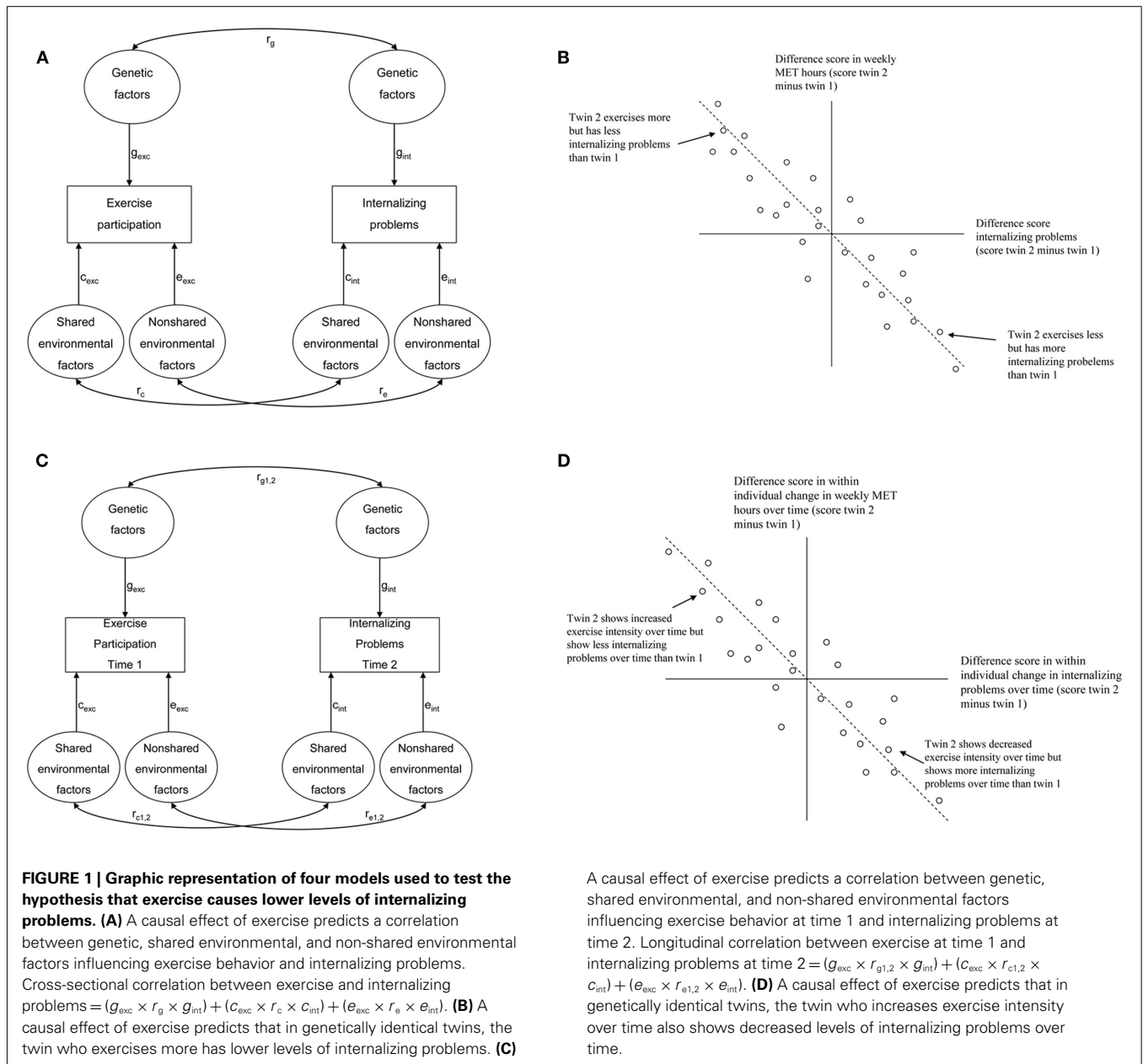
factors influencing exercise behavior, internalizing problems, and SWB. This can be tested in a bivariate genetic model by computing genetic and environmental correlations between exercise behavior at baseline and internalizing problems and SWB at a successive time point. If genetic and environmental correlations between exercise behavior at baseline and internalizing problems or SWB at follow-up are both significant this would be in line with the causal hypothesis (see **Figure 1C**). If only the genetic correlation is significant this is more compatible with the hypothesis of common genetic factors underlying exercise behavior, internalizing problems and SWB.

To extend the procedure by de Moor et al. (2008), a fourth prediction was tested. Under the causal hypothesis, monozygotic intrapair differences in within-individual changes in exercise behavior over time should be associated with parallel intrapair differences in within-individual changes in internalizing problems and SWB over time. In other words, the twin showing the largest increase in frequency and intensity of exercise behavior over time should display larger decreases in internalizing problems and increases in SWB over time than the co-twin showing a lower increase or even a decrease in frequency and intensity of exercise behavior over time (see **Figure 1D**). This would be compatible with the causal hypothesis and it excludes confounding by genetic factors that would independently cause increases in exercise behavior as well as wellbeing over time.

MATERIALS AND METHODS
SUBJECTS
Participants were registered with the Netherlands Twin Registry (NTR), established by the Department of Biological Psychology at the VU University in Amsterdam (Boomsma et al., 2006; Bartels et al., 2007). The large majority of twins had been registered with the NTR as newborns and they were followed throughout childhood. At the time twins were 13 years old, their parents were asked for consent to send their children a survey. If their parents consented, twins and their non-twin-siblings received an online or a paper and pencil self-report survey when they were 14, 16, and 18 years. The survey contained items about behavior, exercise, lifestyle, and wellbeing. In this way, longitudinal data on exercise behavior, internalizing problems, and SWB were collected on three different occasions during adolescence. Data collection started in 2005 and currently data have been collected on two occasions (14 and 16 or 16 and 18). Since our data collection is an ongoing process, data on three different occasions during adolescence will become available in the coming years. When twins and siblings did not return the survey on time they were contacted by mail for a first reminder and next they were contacted by phone for a second reminder. For a detailed description of the adolescent survey collection in the NTR see Bartels et al. (2011). A total of 3645 families with twins born between 1986 and 1994 participated in this ongoing study at least once so far. The overall family response rate is 56%. Because the response rate of the survey collection was moderate, non-response analyses were conducted in which responding and non-responding families were compared on a number of different characteristics, such as socioeconomic status (SES), parental educational level, emotional and behavioral problems of the twins during childhood, and educational achievement

FIGURE 1 | Graphic representation of four models used to test the hypothesis that exercise causes lower levels of internalizing problems. (A) A causal effect of exercise predicts a correlation between genetic, shared environmental, and non-shared environmental factors influencing exercise behavior and internalizing problems. Cross-sectional correlation between exercise and internalizing problems $= (g_{exc} \times r_g \times g_{int}) + (c_{exc} \times r_c \times c_{int}) + (e_{exc} \times r_e \times e_{int})$. (B) A causal effect of exercise predicts that in genetically identical twins, the twin who exercises more has lower levels of internalizing problems. (C)

A causal effect of exercise predicts a correlation between genetic, shared environmental, and non-shared environmental factors influencing exercise behavior at time 1 and internalizing problems at time 2. Longitudinal correlation between exercise at time 1 and internalizing problems at time 2 $= (g_{exc} \times r_{g1,2} \times g_{int}) + (c_{exc} \times r_{c1,2} \times c_{int}) + (e_{exc} \times r_{e1,2} \times e_{int})$. (D) A causal effect of exercise predicts that in genetically identical twins, the twin who increases exercise intensity over time also shows decreased levels of internalizing problems over time.

of the twins. Observed differences on these characteristics were relatively small, i.e., responding and non-responding families were fairly comparable. In other words, response bias is not likely to constitute a serious threat to the validity of the results in the present study.

Twins and siblings with an illness or handicap interfering with their daily lives were not selected for the present study. Information on this came from parental reports from questionnaires in previous years, and from the parental consent form on which parents could indicate if a twin or a sibling was not able to participate in questionnaire research. Data from one additional sibling per family were included. From families with more than one additional participating sibling, we selected the sibling closest in age to the twin. This resulted in exclusion of 157 siblings from

the analyses. Maximum age difference between twins and siblings was 5.7 years (Mean age difference = 2.4 years, SD = 1.0). The total sample consisted of 6317 twins (44% male) and 1180 non-twin-siblings (44% male), from 3486 families, for whom data were available on at least one time point. For the cross-sectional analyses, data from the measurement occasion at the youngest age were selected. The age of the twins in the cross-sectional dataset ranged between 13.2 and 19.9 years, with a mean age of 15.7 years (SD = 1.4) and the age of the non-twin-siblings ranged between 13.0 and 20.0 with a mean age of 16.9 years (SD = 1.9). The cross-sectional dataset consisted of 528 MZ male, 455 DZ male, 778 MZ female, 559 DZ female, 1085 DZ opposite-sex twin pairs, 519 brothers, and 661 sisters. Of these, there were 2912 (86%) complete and 493 (14%) incomplete twin pairs. There were 81

families for which only non-twin-sibling data were available. For 802 (34.6%) of the same-sex twin pairs zygosity was determined based on blood group or DNA typing. Zygosity for the remaining same-sex twin pairs was determined by longitudinally collected questionnaire items about physical similarities and confusion by family members and strangers. These items allow accurate determination of zygosity in 93% of same-sex twin pairs (Rietveld et al., 2000).

Longitudinal analyses were based on a subsample of 1658 twins (38% male) and 150 non-twin-siblings (39% male), from 979 families for whom data on exercise behavior, internalizing problems, and SWB were available at two time points (at age 14 and 16 or 16 and 18). The longitudinal dataset consisted of 135 MZ male, 91 DZ male, 259 MZ female, 179 DZ female, and 289 DZ opposite-sex twin pairs, 58 brothers, and 92 sisters. Of these, there were 705 (74%) complete and 248 (26%) incomplete twin pairs. There were 26 families for which longitudinal data were available for non-twin-siblings only.

MEASURES

Participants were asked to indicate what type(s) of regular leisure exercise they were involved in at the time of assessment. A list of 21 common individual (including for example fitness, jogging) and team-based exercise activities (soccer, field hockey) was provided plus five open entries for less common activities. For each exercise activity endorsed, the participants further reported on the number of months per year, weekly frequency, and average duration of the activity. Ainsworth's Compendium of physical activity (Ainsworth et al., 2000) was used to assign a MET score (Metabolic equivalent) to each exercise activity, reflecting its energy expenditure as a multiple of the basal energy expenditure (approximately 1 kcal/kg/h) in an average subject engaged in that activity.

For each participant a total weekly MET score was computed across all exercise activities by summing the products of the number of hours spent weekly on each exercise activity and its MET score. Activities were only scored if that participant had engaged in them for at least 3 months during the past year. Exercise during physical education classes at school was not included in the weekly MET score, since this is compulsory at all schools in the Netherlands. Thus the dependent variable reflects leisure time exercise behavior only. The distribution of the total weekly MET score was skewed (distribution, 2.58) and demonstrated kurtosis (distribution, 16.06). Therefore, for the bivariate genetic models, log-transformation was applied to the weekly MET score which significantly reduced skewness and kurtosis to −0.51 and −1.12, respectively. For computation of the difference scores, we used untransformed weekly MET scores, because the difference scores were normally distributed.

Self-reported internalizing problems were collected with the Youth Self-Report (YSR; Achenbach and Rescorla, 2001). The YSR was developed for adolescents to rate behavioral and emotional problems. The questionnaire consisted of 112 items from the 2001 version supplemented with six items from older versions of the YSR. The items had to be answered on a three-point scale based on the occurrence of the behavior during the past 6 months: 0 if the item was not true, 1 if the item was somewhat or sometimes true, and 2 if the item was often or very true. We used a Dutch

translation of the YSR which shows good psychometrical properties (Verhulst et al., 1997). The internalizing scale consists of the subscales Anxious/Depressed, Somatic Complaints, and Withdrawn and has 31 items. Example items are "I am too shy" and "I feel lonely." Scores on the individual items were summed to get an overall score of internalizing problems which could range from 0 to 62 with high scores indicating high levels of internalizing problems. Internal consistency of the scale was good with a Chronbach's Alpha of 0.88.

Three measures of self-reported evaluations of SWB were used in the present study: (1) satisfaction with life (SAT) was assessed with the Satisfaction with Life Scale (Diener et al., 1985). The scale consists of five items which had to be answered on a seven-point scale ranging from 1 = "strongly disagree" to 7 = "strongly agree." Example items are "My life is going more or less as I wished" and "I am satisfied with my life." Internal consistency of the scale was good with a Chronbach's Alpha of 0.86. (2) Subjective happiness (HAP) was assessed with the Subjective Happiness Scale (Lyubomirsky and Lepper, 1999). The scale consists of four items which had to be answered on seven-point scale ranging from 1 = "strongly disagree" to 7 = "strongly agree." Example items are "On the whole I am a happy person" and "On the whole, I am not very happy." Internal consistency of the scale was good with a Chronbach's Alpha of 0.84. (3) Quality of life in general (QLg) as assessed with the Cantril Ladder (Cantril, 1965). The ladder has 10 steps: the top indicated the best possible life, and the bottom the worst possible life. Participants had to indicate the step of the ladder at which they place their lives in general.

A genetic factor analysis of SAT, HAP, and QLg demonstrated that these measures assess the same genetic construct (Bartels and Boomsma, 2009). Therefore measures of SAT, HAP, and QLg were summarized into a single score for SWB. For all participants, a latent factor score for SWB was estimated in Mplus (Muthén and Muthén, 2010). Phenotypic correlations between SAT, HAP, and QLg ranged between 0.54 and 0.74. Correlations between the resulting latent factor scores and SAT, HAP, and QLg ranged between 0.70 and 0.95, and the factor loadings were estimated at 0.89, 0.84, and 0.65 respectively, indicating that the latent factor score is a good representation of the three observed measures of SWB.

STATISTICAL ANALYSES

Causal effects of exercise on internalizing problems and SWB were tested in four different models. The first two tests were based on cross-sectional data. In the first test it was assessed whether the genetic (r_g), shared environmental (r_c), or non-shared environmental (r_e) correlation between exercise behavior and internalizing problems and SWB were significantly different from zero. If the genetic (r_g), shared environmental (r_c), or non-shared environmental (r_e) correlation between exercise behavior and internalizing problems, and between exercise behavior and SWB is not significantly different from zero, a causal effect is not likely to be the source of the association. It is important to note that the significance of r_c can only be tested if individual differences in exercise behavior, internalizing problems, and SWB are accounted for by shared environmental factors. If, for example, variation in SWB is not accounted for by shared environmental

factors there will be no r_c between exercise behavior and SWB. In this case, the causal hypothesis will be supported if r_g and r_e are both significant, whereas a causal effect is not likely to be the source of the association if r_g or r_e is not significantly different from zero.

Using maximum-likelihood based structural equation modeling in Mx (Neale et al., 2006), a bivariate genetic model was fitted to the data in which the genetic architecture of the covariation between exercise behavior and internalizing problems and between exercise behavior and SWB was specified and r_g, r_c, and r_e were computed (see **Figure 1A**). The amount of variation in a trait under study and the covariation between traits due to additive genetic (A), shared environmental (C), and non-shared environmental factors (E) can be estimated by considering the different levels of genetic relatedness between monozygotic (MZ) and dizygotic (DZ) twin pairs and non-twin-siblings. MZ twin pairs are genetically identical, whereas DZ twin pairs and non-twin-siblings share on average 50% of their segregating genetic material. Shared environmental effects refer to environmental factors shared by members of the same family. Variation that is not due to genetic and shared environmental effects is attributed to environmental effects which are not shared by family members. The non-shared environmental variance component also includes measurement error variance. Based on the amount of variance in and covariance between traits due to A, C, and E, r_g, r_c, and r_e can be computed. r_g, r_c, and r_e were allowed to differ between boys and girls. Sex differences in the magnitude of r_g, r_c, and r_e were assessed by testing whether constraining r_g, r_c, and r_e for boys and girls to be equal resulted in a significant deterioration of model fit. The hypothesis of a causal effect was assessed by testing whether constraining r_g, r_c, and r_e to zero one by one would lead to a significant deterioration of model fit.

The fit of the different models was compared by means of the log-likelihood ratio test (LRT). The difference in minus two times the log-likelihood ($-2LL$) between two nested models has a χ^2 distribution with the degrees of freedom (df) equaling the difference in df between the two models. If a p-value higher than 0.05 was obtained from the χ^2-test the fit of the constrained model was not significantly worse than the fit of the more complex model. In this case, the constrained model was kept as the most parsimonious and best fitting model.

In the second test of the causal hypothesis we assessed whether MZ twin intrapair differences in levels of exercise participation are associated with intrapair differences in internalizing problems and SWB. Significant associations between intrapair differences in exercise behavior and internalizing problems, and exercise behavior and SWB would be compatible with the causal hypothesis, whereas non-significant associations would falsify this hypothesis. Differences of the measures of exercise behavior, internalizing problems, and SWB of an MZ twin and his or her co-twin were computed. The resulting intrapair difference scores of exercise behavior, internalizing problems, and SWB were normally distributed. Using maximum-likelihood estimation in Mplus, the MZ intrapair differences in internalizing problems and SWB were regressed on the difference in weekly MET hours.

The third and fourth tests of the causal hypothesis were based on the longitudinal dataset. Longitudinal correlations between exercise behavior and internalizing problems and SWB for a 2-year time interval were computed in Mx. If exercise participation causes fewer internalizing problems and increased levels of SWB, there should be a significant longitudinal correlation between exercise behavior at baseline and subsequent internalizing problems and SWB. This significant longitudinal correlation does not necessarily provide evidence for causality if common genetic factors drive the association between exercise behavior at baseline and subsequent internalizing problems and SWB. The genetic information in our sample was again used to more robustly test causality by fitting a longitudinal bivariate genetic model in Mx (see **Figure 1C**). If the effect of exercise behavior on internalizing problems and SWB is causal, $r_{g1,2}$, $r_{c1,2}$, and $r_{e1,2}$ should be significant. $r_{g1,2}$, $r_{c1,2}$, and $r_{e1,2}$ represent the genetic, shared environmental, and non-shared environmental correlation between the measurement of exercise behavior at baseline and successive measurement of internalizing problems and SWB with a 2-year interval.

As with the cross-sectional bivariate genetic model, the genetic architecture of the covariation between exercise behavior at baseline and internalizing problems and SWB at follow-up was specified, and $r_{g1,2}$, $r_{c1,2}$, and $r_{e1,2}$ were estimated. Sex differences in the magnitude of $r_{g1,2}$, $r_{c1,2}$, and $r_{e1,2}$ were assessed by testing whether constraining $r_{g1,2}$, $r_{c1,2}$, and $r_{e1,2}$ for boys and girls to be equal resulted in a significant deterioration of model fit. The hypothesis of a causal effect was assessed by testing whether constraining $r_{g1,2}$, $r_{c1,2}$, and $r_{e1,2}$ to zero would lead to a significant deterioration of model fit. It is important to note that the significance of $r_{c1,2}$ can only be tested if individual differences in exercise behavior, internalizing problems, and SWB are accounted for by shared environmental factors. If, for example, variation in SWB is not accounted for by shared environmental factors there will be no $r_{c1,2}$ between exercise behavior and SWB. In this case, the causal hypothesis will be supported if $r_{g1,2}$ and $r_{e1,2}$ are both significant, whereas a causal effect is not likely to be the source of the association if $r_{g1,2}$ or $r_{e1,2}$ is not significantly different from zero.

In the fourth test of the causal hypothesis we assessed whether MZ twin intrapair differences in individuals' changes in exercise behavior over time are associated with intrapair differences in individuals' changes in internalizing problems and SWB over time (see **Figure 1D**). Significant associations between these intrapair differences would be compatible with the causal hypothesis, whereas non-significant associations would falsify this hypothesis. Within-individual differences between baseline and follow-up scores were computed for the measures of exercise behavior, internalizing problems, and SWB. Subsequently, differences between an MZ twin and his or her co-twin were computed for the within-individual changes in exercise behavior, internalizing problems, and SWB over time. The resulting MZ intrapair difference scores were normally distributed. Using maximum-likelihood estimation in MPlus, the MZ intrapair differences in within-individual changes in internalizing problems and SWB over time were regressed on the MZ intrapair differences of within-individual changes in exercise behavior over time.

RESULTS

In **Table 1**, means and variances of exercise behavior, internalizing problems, and SWB (uncorrected for age) are presented as

a function of sex. Significant effects of age were found on the means of exercise behavior [χ^2 (1) = 43.28, $p \leq 0.01$; standardized regression coefficient $\beta = -0.08$], internalizing problems [χ^2 (1) = 7.83, $p \leq 0.01$; standardized regression coefficient $\beta = 0.03$], and SWB [χ^2 (1) = 26.74, $p \leq 0.01$; standardized regression coefficient $\beta = -0.06$], indicating that mean scores for exercise behavior and SWB were higher for younger participants, whereas the mean score for internalizing problems was higher for older participants. Means of exercise behavior [χ^2 (1) = 149.75, $p \leq 0.01$] and SWB [χ^2 (1) = 355.84, $p \leq 0.01$] were significantly higher for boys, whereas the mean of internalizing problems [χ^2 (1) = 18.69, $p \leq 0.01$] was significantly higher for girls. Variance in exercise behavior [χ^2 (1) = 4.13, $p = 0.04$] was significantly larger for boys, whereas variance in internalizing problems [χ^2 (1) = 251.31, $p \leq 0.01$] and SWB [χ^2 (1) = 23.78, $p \leq 0.01$] were significantly higher for girls.

Twin and twin-sibling correlations for exercise behavior, internalizing problems and SWB are presented in the left part of **Table 2**. The DZ twin correlations for the traits were equal to twin-sibling correlations (all p-values > 0.05). The MZ twin correlations were significantly larger than the DZ twin and twin-sibling correlations (all p-values < 0.01), indicating that genetic influences accounted for individual differences in exercise behavior, internalizing problems and SWB. For girls, DZ twin, and twin-sibling correlations for exercise behavior and internalizing problems were higher than half the MZ twin correlations, suggesting that shared environmental factors play a role in explaining variation in exercise behavior and internalizing problems. For exercise behavior and internalizing problems in boys and SWB in boys and girls, DZ twin, and twin-sibling correlations were about half the MZ twin correlations, suggesting no shared environmental effects.

CROSS-SECTIONAL ASSOCIATIONS

The phenotypic correlation between exercise behavior and internalizing problems was −0.13 [95% CI: −0.14 to −0.10; χ^2 (1) = 93.81, $p \leq 0.01$] and between exercise behavior and SWB it was 0.12 [95% CI: 0.09 to −0.15; χ^2 (1) = 91.90, $p \leq 0.01$]. The phenotypic correlations were equal for boys and girls (all p-values > 0.05). This indicates that increased levels of exercise behavior are cross-sectionally associated with decreased levels of internalizing problems and increased levels of SWB. The phenotypic correlation between internalizing problems and SWB was −0.51 [95% CI: −0.53 to −0.49; χ^2 (1) = 1766.78, $p \leq 0.01$]. Cross-twin cross-trait correlations are presented in the right part of **Table 2**. The cross-twin cross-trait correlations in MZ and DZ twin pairs, and twin-sibling pairs were significant (all p-values < 0.05). This indicates that exercise behavior in one twin is predictive of internalizing problems and SWB in his or her co-twin or non-twin-sibling. The DZ cross-twin cross-trait correlations were equal to twin-sibling cross-trait correlations (p-values exceeded a 5% significance level). The MZ cross-twin cross-trait correlations were significantly larger than the DZ cross-twin cross-trait and the twin-sibling cross-trait correlations (all p-values < 0.05), indicating that genetic influences partly account for the covariation between exercise behavior and internalizing problems, and between exercise behavior and SWB. MZ and DZ cross-twin cross-trait correlations were equal for boys and girls (all p-values > 0.05).

Bivariate genetic model

Bivariate genetic model fitting results are presented in **Table 3**. In the upper part of the table the results of the bivariate genetic model for exercise behavior and internalizing problems are presented. r_c was not modeled for boys, because shared environmental influences did not contribute to variation in exercise behavior and internalizing problems in boys (Lamb et al., 2010; van der Aa et al., 2010a; Bartels et al., 2011) In model 2 the genetic (r_g) and non-shared environmental correlation (r_e) were constrained to be equal for boys and girls which did not result in a significant deterioration of model. Subsequent models were therefore compared to this model. Models 3–5 tested whether constraining r_g, r_c for girls, and r_e to zero would lead to a significant deterioration of model fit. r_g was statistically significant for boys and, and r_c was statistically significant for girls, whereas r_e was not statistically significant. r_g was −0.19 (95% CI: −0.26 to −0.15) and r_c for girls

Table 1 | Maximum-likelihood estimates of means and variances of exercise behavior, internalizing problems, and SWB as a function of sex.

		M	Var
Exercise behavior (weekly MET score)	Boys	26.8	664.8
	Girls	18.4	521.9
Internalizing problems (scale score)	Boys	6.8	31.1
	Girls	9.6	55.3
SWB (factor score)	Boys	0.06	0.79
	Girls	−0.04	0.93

Table 2 | Twin and twin-sibling correlations for exercise behavior, internalizing problems and SWB, and the cross-twin cross-trait and cross-sibling cross-trait correlations of exercise behavior, internalizing problems, and subjective wellbeing.

	Exercise	Internalizing	SWB	Exercise–internalizing	Exercise–SWB
MZM	0.69 (0.65–0.73)	0.63 (0.56–0.68)	0.41 (0.33–0.48)	−0.11 (−0.16 to −0.06)	0.10 (0.06–0.15)
DZM/sibling	0.32 (0.26–0.38)	0.34 (0.25–0.43)	0.11 (0.03–0.19)	−0.07 (−0.13 to −0.01)	0.07 (0.02–0.12)
MZF	0.72 (0.69–0.75)	0.50 (0.45–0.55)	0.47 (0.41–0.52)	−0.12 (−0.16 to −0.09)	0.14 (0.11–0.18)
DZF/sibling	0.42 (0.36–0.47)	0.30 (0.25–0.35)	0.24 (0.18–0.30)	−0.08 (−0.12 to −0.04)	0.11 (0.06–0.15)
DOS/sibling	0.28 (0.24–0.33)	0.24 (0.19–0.28)	0.20 (0.15–0.24)	−0.05 (−0.09 to −0.02)	0.08 (0.05–0.12)

MZ, monozygotic twin pair; DZ, dizygotic twin pair; M, male; F, female; DOS, opposite-sex twin pair.

Table 3 | Bivariate model fitting results for exercise behavior and internalizing problems and SWB.

Model	vs.	−2LL	df	χ²	df	p
EXERCISE BEHAVIOR–INTERNALIZING PROBLEMS						
1. r_g, r_e boys; r_g, r_c, r_e girls: sex differences	–	58490.629	14427	–	–	–
2. r_g, r_e boys; r_g, r_c, r_e girls: no sex differences	1	58492.491	14429	1.86	2	0.39
3. $r_g = 0$	2	58514.756	14430	22.26	1	<0.01
4. r_c girls $= 0$	2	58503.387	14430	10.90	1	<0.01
5. $r_e = 0$	2	58493.414	14430	0.92	1	0.34
EXERCISE BEHAVIOR–SWB						
1. r_g, r_e: sex differences	–	32001.879	14602	–	–	–
2. r_g, r_e: no sex differences	1	32009.817	14604	7.94	2	0.02
3a. r_g boys $= 0$	1	32017.844	14603	15.97	1	<0.01
3b. r_g girls $= 0$	1	32064.555	14603	62.68	1	<0.01
4a. r_e boys $= 0$	1	32002.042	14603	0.16	1	0.69
4b. r_e girls $= 0$	1	32006.982	14603	5.10	1	0.02

r_g, genetic correlation; r_c, shared environmental correlation; r_e, non-shared environmental correlation; vs., versus; −2LL, −2 log-likelihood; df, degrees of freedom; χ², chi-square test statistic; df, degrees of freedom of χ²-test; p, p-value.

was −0.45 (95% CI: −1.0 to −0.19). These results suggest that the association between exercise behavior and internalizing problems in boys seems to be mainly due to genetic factors which are common (i.e., overlapping) for exercise behavior and internalizing problems. For girls, genetic and shared environmental factors both account for the association. Because r_e was not different from zero the causal hypothesis was not supported in either boys or girls.

In the lower part of **Table 3** the results of the bivariate genetic model for exercise behavior and SWB are presented. For SWB, r_c was not modeled, because shared environmental influences did not contribute to variation in SWB in boys and girls (Bartels and Boomsma, 2009; van der Aa et al., 2010b). In model 2, r_g and r_e were constrained to be equal for boys and girls which resulted in a significant deterioration of model fit. Therefore, r_g and r_e were allowed to differ between boys and girls. Model 3 and 4 tested whether constraining r_g and r_e to zero would lead to a significant deterioration of model fit. r_g was statistically significant in boys and girls and estimated at 0.22 (95% CI: 0.14 to −0.29) and 0.31 (95% CI: 0.23 to −0.39) for boys and girls respectively. For boys, r_e was not significantly different from zero, $p \leq 0.01$ and estimated at 0.02 (95% CI: −0.06 to 0.10). For girls, r_e was significant and close to zero (0.05; 95% CI: 0.01 to −0.09). These results suggested that the association between exercise behavior and SWB in girls and boys is mainly due to genetic pleiotropy.
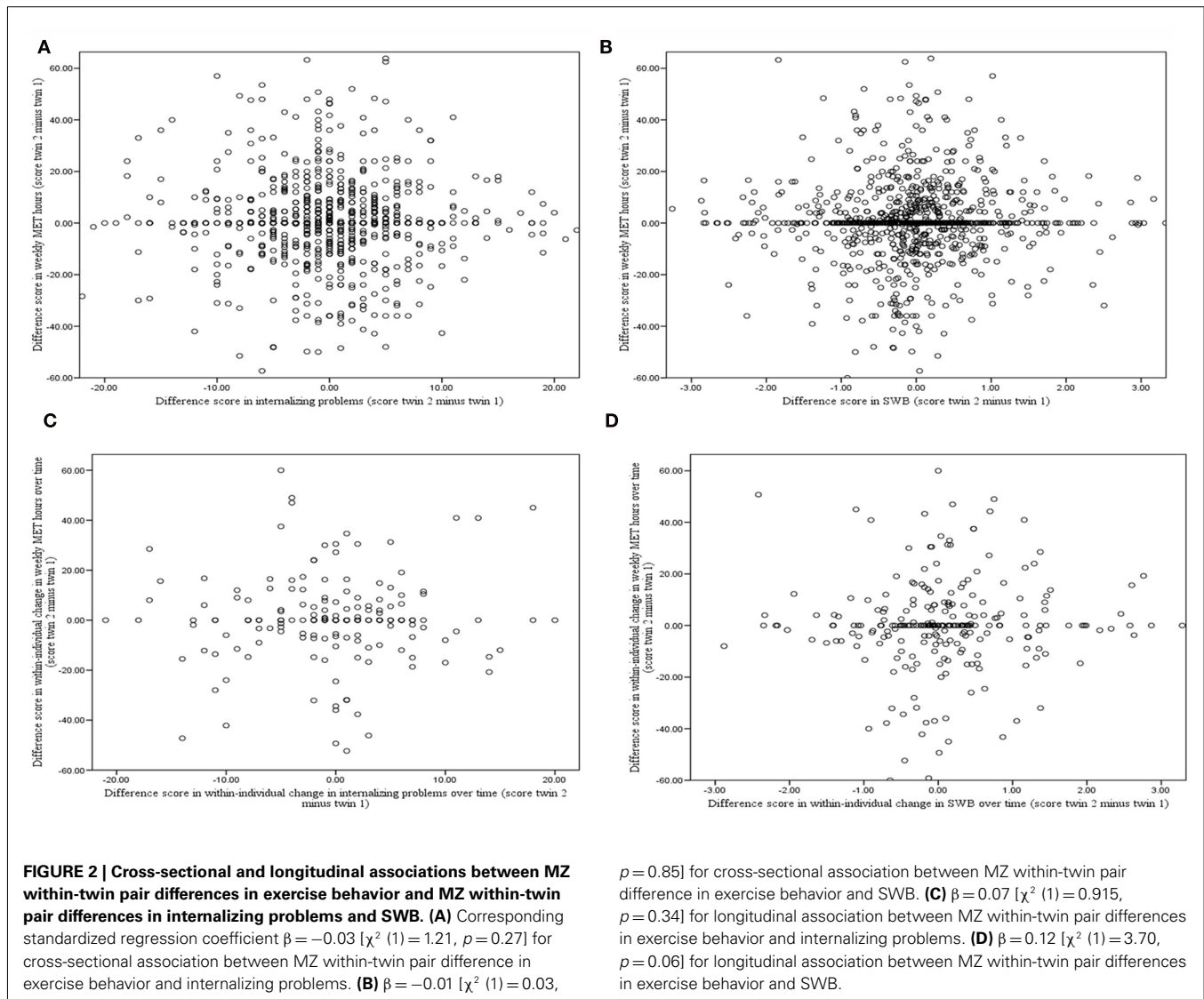
MZ twin intrapair differences model

In the upper part of **Figure 2**, cross-sectional associations between MZ within-twin pair difference scores in exercise behavior and internalizing problems (**Figure 2A**) and SWB (**Figure 2B**) are presented. It can be seen that there were no significant associations between within-twin pair differences in exercise behavior and internalizing problems and SWB. This indicates that in genetically identical twin pairs, a twin who exercises more does not have fewer internalizing problems or higher levels of SWB than his or her co-twin who exercises less. This is not in line

with the causal hypothesis that increased levels of exercise cause decreased levels of internalizing problems and increased levels of SWB.

LONGITUDINAL ASSOCIATIONS

Based on the longitudinal subsample, twin and twin-sibling correlations for exercise behavior, internalizing problems and SWB at the first (T1) and second (T2) measurement occasion are presented in the left part of **Table 4**. The DZ twin correlations for the traits at both measurement occasions were equal to twin-sibling correlations (p-values > 0.05). Twin and twin-sibling correlations on T1 and T2 were largely similar to those presented in **Table 2** which were based on the cross-sectional sample.

The longitudinal phenotypic correlation between exercise behavior and internalizing problems 2 years later was −0.15 [95% CI: −0.21 to −0.08; $\chi^2 (1) = 19.70$, $p \leq 0.01$] and between exercise behavior and SWB it was 0.16 [95% CI: 0.11 to −0.21; $\chi^2 (1) = 37.66$, $p \leq 0.01$] which is comparable with the magnitude of the cross-sectional phenotypic correlations. This means that exercise behavior at baseline is associated with internalizing problems and SWB at follow-up. Longitudinal cross-twin cross-trait correlations between exercise behavior at T1 and internalizing problems and SWB at T2 are presented in the right part of **Table 4**. The DZ cross-twin cross-trait correlations were equal to twin-sibling cross-trait correlations (p-values > 0.05). For boys, the MZ cross-twin cross-trait correlations were significant larger than the DZ cross-twin cross-trait and the twin-sibling cross-trait correlations (p-values < 0.05), indicating that genetic influences account for the covariation between exercise behavior and internalizing problems, and between exercise behavior and SWB. For girls, the MZ cross-twin cross-trait correlations were not significantly different from the DZ cross-twin cross-trait and the twin-sibling cross-trait correlations (p-values > 0.05), indicating that shared environmental influences account for the covariation between exercise behavior and internalizing problems, and between exercise behavior and SWB.

FIGURE 2 | Cross-sectional and longitudinal associations between MZ within-twin pair differences in exercise behavior and MZ within-twin pair differences in internalizing problems and SWB. (A) Corresponding standardized regression coefficient $\beta = -0.03$ [χ^2 (1) = 1.21, $p = 0.27$] for cross-sectional association between MZ within-twin pair difference in exercise behavior and internalizing problems. **(B)** $\beta = -0.01$ [χ^2 (1) = 0.03, $p = 0.85$] for cross-sectional association between MZ within-twin pair difference in exercise behavior and SWB. **(C)** $\beta = 0.07$ [χ^2 (1) = 0.915, $p = 0.34$] for longitudinal association between MZ within-twin pair differences in exercise behavior and internalizing problems. **(D)** $\beta = 0.12$ [χ^2 (1) = 3.70, $p = 0.06$] for longitudinal association between MZ within-twin pair differences in exercise behavior and SWB.

Longitudinal bivariate genetic model

Longitudinal bivariate model fitting results are presented in **Table 5**. In the upper part of the table results of the longitudinal genetic model for exercise behavior and internalizing problems are presented. Model 2 tested whether constraining the genetic ($r_{g1,2}$) and non-shared environmental correlation ($r_{e1,2}$) to be equal for boys and girls led to a significant deterioration of model fit. $r_{g1,2}$ and $r_{e1,2}$ were not significantly different for boys and girls and were therefore constrained to be equal. Models 3–5 tested whether constraining $r_{g1,2}$ and $r_{e1,2}$, for boys and girls, and $r_{c1,2}$ for girls, to zero would lead to a significant deterioration of model fit. $r_{g1,2}$ was -0.27 (95% CI: -0.56 to -0.10) and significantly larger than zero, whereas $r_{c1,2}$ for girls and $r_{e1,2}$ were not different from zero. This suggests that the longitudinal association between exercise behavior and internalizing problems in both sexes is mainly accounted for by genetic factors which are common for exercise behavior and internalizing problems. Because r_e and r_c for girls were not different from zero the causal hypothesis was not supported.

In the lower part of **Table 5** results of the longitudinal genetic model for exercise behavior and SWB are presented. Model 2 tests whether constraining $r_{g1,2}$ and $r_{e1,2}$ to be equal for boys and girls resulted in a significant deterioration of model fit. $r_{g1,2}$ and $r_{e1,2}$ were not significantly different for boys and girls. Models 3 and 4 tested whether constraining $r_{g1,2}$ and $r_{e1,2}$ to zero would lead to a significant deterioration of model fit. $r_{g1,2}$ was statistically significant, whereas $r_{e1,2}$ was not significantly larger than zero. $r_{g1,2}$ was 0.31 (95% CI: 0.19 to -0.43). This suggests that the longitudinal association between exercise behavior and SWB is also best explained by genetic factors which are common (i.e., overlapping) for exercise behavior and SWB. Because r_e was not different from zero the causal hypothesis was not supported.

Longitudinal MZ twin intrapair differences model

In the lower part of **Figure 2**, associations between MZ within-twin pair difference scores in individuals' changes in exercise behavior over time and MZ within-twin pair difference scores in individuals' changes in internalizing problems (**Figure 2C**) and SWB

Table 4 | Twin and twin-sibling correlations for exercise behavior, internalizing problems and SWB at the first (T1) and second (T2) measurement occasion, and the longitudinal cross-twin cross-trait and cross-sibling cross-trait correlations for exercise behavior, internalizing problems and subjective wellbeing.

	Exercise		Internalizing		SWB		Exercise T1–internalizing T2	Exercise T1–SWB T2
	T1	T2	T1	T2	T1	T2		
MZM	0.74	0.66	0.71	0.40	0.28	0.43	−0.13	0.13
	(0.66–0.81)	(0.56–0.74)	(0.60–0.78)	(0.09–0.59)	(0.09–0.44)	(0.25–0.57)	(−0.23 to −0.02)	(0.04–0.22)
DZM/sibling	0.27	0.21	0.27	0.00	0.03	−0.01	0.07	−0.05
	(0.12–0.40)	(0.02–0.38)	(0.08–0.43)	(−0.22 to 0.25)	(−0.15 to 0.22)	(−0.18 to 0.17)	(−0.10 to 0.23)	(−0.19 to 0.09)
MZF	0.81	0.75	0.57	0.44	0.52	0.36	−0.14	0.16
	(0.76–0.84)	(0.69–0.79)	(0.48–0.65)	(0.32–0.55)	(0.41–0.60)	(0.24–0.46)	(−0.22 to −0.07)	(0.10–0.22)
DZF/sibling	0.42	0.45	0.32	0.39	0.22	0.12	−0.07	0.19
	(0.33–0.50)	(0.33–0.55)	(0.23–0.40)	(0.24–0.51)	(0.12–0.30)	(−0.06 to 0.28)	(−0.17 to 0.03)	(0.09–0.28)
DOS/sibling	0.36	0.24	0.24	0.20	0.15	0.15	−0.08	0.09
	(0.28–0.43)	(0.14–0.34)	(0.14–0.33)	(0.00–0.36)	(0.06–0.23)	(0.02–0.27)	(−0.18 to 0.02)	(0.01–0.18)

T1, first measurement occasion; T2, second measurement occasion; MZ, monozygotic twin pair; DZ, dizygotic twin pair; M, male; F, female; DOS, opposite-sex twin pair.

Table 5 | Longitudinal bivariate model fitting results for exercise behavior and internalizing problems and SWB.

Model	vs.	−2LL	df	χ^2	df	p
EXERCISE BEHAVIOR–INTERNALIZING PROBLEMS						
1. $r_{g1,2}$, $r_{e1,2}$ boys; $r_{g1,2}$, $r_{c1,2}$, $r_{e1,2}$ girls: sex differences	–	9583.709	2838	–	–	–
2. $r_{g1,2}$, $r_{e1,2}$ boys; $r_{g1,2}$, $r_{c1,2}$, $r_{e1,2}$ girls: no sex differences	1	9584.841	2840	1.13	2	0.57
3. $r_{g1,2} = 0$	2	9588.744	2841	3.90	1	0.05
4. $r_{c1,2}$ girls $= 0$	2	9586.936	2841	2.10	1	0.15
5. $r_{e1,2} = 0$	2	9585.028	2841	0.19	1	0.67
EXERCISE BEHAVIOR–SWB						
1. $r_{g1,2}$, $r_{e1,2}$: sex differences	–	7237.608	3455	–	–	–
2. $r_{g1,2}$, $r_{e1,2}$: no sex differences	1	7239.709	3457	2.10	2	0.35
3. $r_{g1,2} = 0$	2	7265.019	3458	25.31	1	<0.01
4. $r_{e1,2} = 0$	2	7239.716	3458	0.01	1	0.93
5. $r_{g1,2} = r_{g1,2,exp}$	2	7245.787	3458	6.08	1	0.01
6. $r_{e1,2} = r_{e1,2,exp}$	2	7248.324	3458	8.62	1	<0.01

$r_{g1,2}$, genetic correlation; $r_{c1,2}$, shared environmental correlation; $r_{e1,2}$, non-shared environmental correlation; vs., versus; −2LL, −2 log-likelihood; df, degrees of freedom; χ^2, chi-square test statistic; df, degrees of freedom of χ^2-test; p, p-value.

(**Figure 2D**) over time are presented. It can be seen that there were no significant associations between within-twin pair differences in exercise behavior and internalizing problems and SWB. This indicates that in genetically identical twin pairs, a twin showing increased frequency and intensity of exercise behavior over time does not show lower levels of internalizing problems or higher levels of SWB over time than his or her co-twin with decreasing levels of exercise behavior over time. This is again not in line with the causal hypothesis that increased levels of exercise cause decreased levels of internalizing problems and increased levels of SWB.

DISCUSSION

The present study investigated, in a large sample of Dutch adolescent twins and their non-twin-siblings, whether exercise behavior causally influences internalizing problems and SWB or whether the association between these traits reflects the effects of underlying genetic factors. In keeping with population-based studies investigating adult samples (Camacho et al., 1991; Norris et al., 1992; de Moor et al., 2006, 2008; Wise et al., 2006; Stubbe et al., 2007), we found that exercise behavior was cross-sectionally and longitudinally associated with fewer internalizing problems and increased levels of SWB across the total sample of adolescents. However, within genetically identical twins, a twin who exercised more did not have fewer internalizing problems or higher levels of SWB than his or her less exercising co-twin and, in addition, a twin showing increased frequency and intensity of exercise behavior over time did not show lower levels of internalizing problems or higher levels of SWB over time than his or her co-twin with

unchanged or decreased levels of exercise behavior. The cross-sectional and longitudinal associations between exercise behavior and internalizing problems and SWB were entirely accounted for by genetic factors in boys and genetic and shared environmental factors in girls, whereas the unique environmental factors causing them to participate in exercise did not cause lower levels of internalizing problems and increased levels of SWB, currently, or 2 years later.

Taken together, these results do not support the hypothesis that increased levels of psychological wellbeing found in adolescent exercisers reflect a causal effect of exercise. Instead, the results are compatible with the alternative hypothesis that the association between exercise behavior and psychological wellbeing in adolescence is mainly due to an underlying factor. Pleiotropic genes seem to be the major component of this underlying factor although shared environmental influences also contribute to both exercise behavior and internalizing behavior in girls. In adults, pleiotropic genetic factors have also been found to explain the association between exercise and depressive and anxious symptoms (de Moor et al., 2008), as well as between exercise and SWB (Stubbe et al., 2007).

The lack of support for the hypothesis that increased levels of psychological wellbeing found in exercisers reflect a causal effect of exercise needs to be reconciled with findings from training studies that reported improved mood or reductions in coping behavior, depression, and anxiety after a program of aerobic exercise training in comparison to control manipulations (Steptoe et al., 1989; Brosse et al., 2002; Barbour and Blumenthal, 2005). Although others have failed to replicate these training effects (King et al., 1989; de Geus et al., 1993) and critical reviews on this topic express only cautious optimism about the causal role of exercise (Lawlor and Hopker, 2001; Dunn et al., 2005), recent well-designed studies in depression patients report beneficial psychological effects of exercise that match or even exceed those of pharmacological treatment (Blumenthal et al., 2007). It should be recognized, however, that these studies examined the effects of prescribed and externally monitored exercise treatments in select subgroups and that methodological problems cling to even the most well-designed randomized controlled trial due to the strong societal beliefs about the efficacy of this intervention. Furthermore, there have been no randomized controlled training trials in adolescents that applied the same rigor as the trials in adults. Taken our evidence for genetic pleiotropy and the pitfalls in existing randomized controlled trials for adults, i.e., assuming that regular exercise may prevent depression in a non-clinical adolescent sample, because it has been successfully used as a therapy in subsets of clinically depressed adult patients, is still a big leap of faith.

In spite of the above, we do not rule out the possibility that exercise might have a causal effect on psychological wellbeing in specific subgroups of adolescents. In fact, a model has been proposed that is compatible with genetic pleiotropy and still allows exercise to causally increase wellbeing in specific subgroups. To maintain regular exercise participation, appetitive effects of exercise need to outweigh the aversive effects. Individuals for whom the aversive effects of exercise are stronger than the rewarding effects will eventually cease their exercise participation. In contrast, individuals for whom the appetitive effects are stronger than the aversive effects will maintain exercise participation and become regular exercisers. It has indeed been shown that regular exercisers report greater acute exercise-induced mood enhancement than non-exercisers (Gauvin, 1990; Hoffman and Hoffman, 2008). These positive mood experiences after exercise activities may be a positive reward and therefore an important component of the higher wellbeing reported by regular exercisers. Genetic variants tipping the balance between appetitive and aversive effects may be a major source of individual differences in the effects of exercise on these acute mood responses. These genetic effects on the acute mood response to exercise may not only keep adolescents maintaining regular exercise participation. The repeated exercise-induced increases in mood may influence their overall psychological wellbeing. It is important to note that the causal effect of exercise in this proposed model would be limited to a genetically sensitive subgroup and would not apply to the total adolescent population.

In addition, there is also a social–psychological mechanism through which the association between exercise behavior and psychological wellbeing may be explained by genetic pleiotropy (de Geus and de Moor, 2008). The basic idea is that people will seek out the activities that they are good in. This is particularly true during adolescence, because being good in sports is an important source of self-esteem for adolescents and the athletic role model is continuously reinforced by the media (Field et al., 1999; Pope et al., 2001). Therefore, genes coding for exercise ability (endurance, strength, flexibility, motor coordination) may well become genes for adolescent exercise behavior. These genetic factors may further influence psychological wellbeing because high exercise ability will coincide with the experience of strong feelings of competence and mastery as well as positive feedback from peers, leading to an increased sense of self-esteem and psychological wellbeing. In short, genetic variants influencing exercise ability will act to maintain exercise behavior and to increase psychological wellbeing, but again are limited to a selected group of adolescents, i.e., those with high exercise ability.

Apart from the gene-by-exercise interactive effects sketched above, we cannot rule out the presence of more complex mechanisms, like reverse or reciprocal causality, or combinations of causal, reverse causal, and pleiotropic mechanisms. These complex mechanisms should be assessed in future studies making use of more complex causation models and larger sample sizes. Nonetheless, our results already have implications for the prevention of internalizing problems and decreasing levels of SWB during adolescence and the maintenance of exercise behavior at a sufficient level throughout adolescence and during the transition to adulthood. Current intervention strategies are mainly based on the idea that the associations between exercise and psychological wellbeing reflect causal effects of exercise and that exercise interventions exert beneficial effects on all participants. While it remains possible that exercise behavior can prevent some adolescents from developing internalizing problems, our results show that those with a disposition toward internalizing problems also have a higher chance to be a non-exerciser because of their genetic make-up. Approaching these adolescents at risk for internalizing problems with demanding exercise programs may be doomed to

fail and may even amplify feelings of low self-esteem and consequently internalizing problems. Instead, behavioral therapy aimed to improve psychological wellbeing may need to coincide with or even precede intervention on exercise behavior.

In addition, we cannot fully rule out the effects of genetic rater bias, i.e., the presence of a polygenic factor that predisposes to reporting higher amounts of regular exercise and reporting higher SWB. However, the findings of our longitudinal analyses show that within-individual change in exercise behavior (or reporting of it) over time did not lead to change in wellbeing (or reporting of it) which argues against a significant effect of rater bias.

It is important to note that DZ twins were as similar as non-twin-siblings with regard to exercise behavior, internalizing problems and SWB. In addition, cross-twin cross-trait correlations were also similar for DZ twins and non-twin-siblings. This implicates that the present results do not only apply to twins, but that they can be generalized to the total adolescent population.

In conclusion, among Dutch adolescents, regular exercise was found to be associated with fewer internalizing problems and higher levels of SWB. The association could not be explained by causal effects of exercise behavior. Instead, the association was largely accounted for by genetic factors influencing a tendency toward exercise as well as a disposition toward low levels of internalizing problems and high levels of SWB.

REFERENCES

Achenbach, T. M., and Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms and Profiles.* Burlington, VT: University of Vermont.

Ainsworth, B. E., Haskell, W. L., Whitt, M. C., Irwin, M. L., Swartz, A. M., Strath, S. J., O'Brien, W. L., Bassett, D. R. Jr., and Schmitz, K. H., Emplaincourt, P. O., Jacobs, D. R. Jr. and Leon, A. S. (2000). Compendium of physical activities: an update of activity codes and MET intensities. *Med. Sci. Sports Exerc.* 32, S498–S504.

Barbour, K. A., and Blumenthal, J. A. (2005). Exercise training and depression in older adults. *Neurobiol. Aging* 26(Suppl. 1), 119–123.

Bartels, M., and Boomsma, D. I. (2009). Born to be happy? The etiology of subjective well-being. *Behav. Genet.* 39, 605–615.

Bartels, M., van Beijsterveldt, C. E., Derks, E. M., Stroet, T. M., Polderman, T. J., Hudziak, J. J., and Boomsma, D. I. (2007). Young Netherlands Twin Register (Y-NTR): a longitudinal multiple informant study of problem behavior. *Twin Res. Hum. Genet.* 10, 3–11.

Bartels, M., van der Aa, N., van Beijsterveldt, C. E., Middeldorp, C. M., and Boomsma, D. I. (2011). Adolescent self-report of emotional and behavioral problems: interactions of genetic factors with sex and age. *J. Can. Acad. Child Adolesc. Psychiatry* 20, 35–52.

Beunen, G., and Thomis, M. (1999). Genetic determinants of sports participation and daily physical activity. *Int. J. Obes. Relat. Metab. Disord.* 23(Suppl. 3), S55–S63.

Blumenthal, J. A., Babyak, M. A., Doraiswamy, P. M., Watkins, L., Hoffman, B. M., Barbour, K. A., Herman, S., Craighead, W. E., Brosse, A. L., Waugh, R., Hinderliter, A., and Sherwood, A. (2007). Exercise and pharmacotherapy in the treatment of major depressive disorder. *Psychosom. Med.* 69, 587–596.

Boomsma, D. I., de Geus, E. J., Vink, J. M., Stubbe, J. H., Distel, M. A., Hottenga, J. J., Posthuma, D., van Beijsterveldt, T. C., Hudziak, J. J., Bartels, M., and Willemsen, G. (2006). Netherlands twin register: from twins to twin families. *Twin Res. Hum. Genet.* 9, 849–857.

Brosse, A. L., Sheets, E. S., Lett, H. S., and Blumenthal, J. A. (2002). Exercise and the treatment of clinical depression in adults: recent findings and future directions. *Sports Med.* 32, 741–760.

Camacho, T. C., Roberts, R. E., Lazarus, N. B., Kaplan, G. A., and Cohen, R. D. (1991). Physical activity and depression: evidence from the Alameda County Study. *Am. J. Epidemiol.* 134, 220–231.

Cantril, H. (1965). *The Pattern of Human Concerns.* New Brunswick, NJ: Rutgers University Press.

Costello, E. J., Egger, H., and Angold, A. (2005). 10-Year research update review: the epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden. *J. Am. Acad. Child Adolesc. Psychiatry* 44, 972–986.

de Geus, E. J., van Doornen, L. J., and Orlebeke, J. F. (1993). Regular exercise and aerobic fitness in relation to psychological make-up and physiological stress reactivity. *Psychosom. Med.* 55, 347–363.

de Geus, E. J. C., and de Moor, M. H. (2008). A genetic perspective on the association between exercise and metal health. *Ment. Health Phys. Act.* 1, 53–51.

de Moor, M. H., Beem, A. L., Stubbe, J. H., Boomsma, D. I., and de Geus, E. J. (2006). Regular exercise, anxiety, depression and personality: a population-based study. *Prev. Med.* 42, 273–279.

de Moor, M. H., Boomsma, D. I., Stubbe, J. H., Willemsen, G., and de Geus, E. J. (2008). Testing causality in the association between regular exercise and symptoms of anxiety and depression. *Arch. Gen. Psychiatry* 65, 897–905.

Diener, E., Emmons, R. A., Larsen, R. J., and Griffin, S. (1985). The satisfaction with life scale. *J. Pers. Assess.* 49, 71–75.

Dunn, A. L., Trivedi, M. H., Kampert, J. B., Clark, C. G., and Chambliss, H. O. (2005). Exercise treatment for depression: efficacy and dose response. *Am. J. Prev. Med.* 28, 1–8.

Field, A. E., Cheung, L., Wolf, A. M., Herzog, D. B., Gortmaker, S. L., and Colditz, G. A. (1999). Exposure to the mass media and weight concerns among girls. *Pediatrics* 103, E36.

Gauvin, L. (1990). An experiential perspective on the motivational features of exercise and lifestyle. *Can. J. Sport Sci.* 15, 51–58.

Hoffman, M. D., and Hoffman, D. R. (2008). Exercisers achieve greater acute exercise-induced mood enhancement than nonexercisers. *Arch. Phys. Med. Rehabil.* 89, 358–363.

King, A. C., Taylor, C. B., Haskell, W. L., and DeBusk, R. F. (1989). Influence of regular aerobic exercise on psychological health: a randomized, controlled trial of healthy middle-aged adults. *Health Psychol.* 8, 305–324.

Lamb, D. J., Middeldorp, C. M., van Beijsterveldt, C. E., Bartels, M., van der, A. N., Polderman, T. J., and Boomsma, D. I. (2010). Heritability of anxious-depressive and withdrawn behavior: age-related changes during adolescence. *J. Am. Acad. Child Adolesc. Psychiatry* 49, 248–255.

Lawlor, D. A., and Hopker, S. W. (2001). The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised controlled trials. *BMJ* 322, 763–767.

Lyubomirsky, S., and Lepper, H. S. (1999). A measure of subjective happiness: preliminary reliability and construct validation. *Soc. Indic. Res.* 46, 137–155.

Maia, J. A., Thomis, M., and Beunen, G. (2002). Genetic factors in physical activity levels: a twin study. *Am. J. Prev. Med.* 23, 87–91.

Muthén, L. K., and Muthén, B. O. (2010). *Mplus User's Guide,* 6th Edn. Los Angeles, CA: Muthén and Muthén.

Neale, M. C., Boker, S. M., Xie, G., and Maes, H. H. (2006). *Mx: Statistical Modeling*, 7th Edn. Richmond, VA: Virginia Commonwealth University.

Norris, R., Carroll, D., and Cochrane, R. (1992). The effects of physical activity and exercise training on psychological stress and well-being in an adolescent population. *J. Psychosom. Res.* 36, 55–65.

Pope, H. G. Jr., Olivardia, R., Borowiecki, J. J. III, and Cohane, G. H. (2001). The growing commercial value of the male body: a longitudinal survey of advertising in women's magazines. *Psychother. Psychosom.* 70, 189–192.

Proctor, C. L., Linley, P. A., and Maltby, J. (2009). Youth life satisfaction; a review of the literature. *J. Happiness Stud.* 10, 583–630.

Rietveld, M. J., van Der Valk, J. C., Bongers, I. L., Stroet, T. M., Slagboom, P. E., and Boomsma, D. I. (2000). Zygosity diagnosis in young twins by parental report. *Twin Res.* 3, 134–141.

Steptoe, A., Edwards, S., Moses, J., and Mathews, A. (1989). The effects of exercise training on mood and perceived coping ability in anxious adults from the general population. *J. Psychosom. Res.* 33, 537–547.

Stubbe, J. H., Boomsma, D. I., and de Geus, E. J. (2005). Sports participation during adolescence: a shift from environmental to genetic factors. *Med. Sci. Sports Exerc.* 37, 563–570.

Stubbe, J. H., de Moor, M. H., Boomsma, D. I., and de Geus, E. J. (2007). The association between exercise participation and well-being: a co-twin study. *Prev. Med.* 44, 148–152.

van Mechelen, W., Twisk, J. W., Post, G. B., Snel, J., and Kemper, H. C. (2000). Physical activity of young people: the Amsterdam Longitudinal Growth and Health Study. *Med. Sci. Sports Exerc.* 32, 1610–1616.

van der Aa, N., de Geus, E. J., van Beijsterveldt, T. C., Boomsma, D. I., and Bartels, M. (2010a). Genetic influences on individual differences in exercise behavior during adolescence. *Int. J. Pediatr.* 2010, 138345.

van der Aa, N., Boomsma, D. I., Rebollo-Mesa, I., Hudziak, J. J., and Bartels, M. (2010b). Moderation of genetic factors by parental divorce in adolescents' evaluations of family functioning and mental well-being. *Twin Res. Hum. Genet.* 13, 143–162.

Verhulst, F. C., van der Ende, J., and Koot, H. M. (1997). *Handleiding voor de Youth Self-Report (YSR) [Manual for the Youth Self-Report (YSR)]*. Rotterdam: Academisch Ziekenhuis Rotterdam/Erasmus Universiteit Rotterdam.

Wise, L. A., Adams-Campbell, L. L., Palmer, J. R., and Rosenberg, L. (2006). Leisure time physical activity in relation to depressive symptoms in the Black Women's Health Study. *Ann. Behav. Med.* 32, 68–76.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.