Summary and Discussion

Summary

The past years have seen a remarkable surge of scientific discoveries in human genetics, and complex-trait genetics specifically. This has been largely facilitated by the development of affordable genome-wide arrays and aided by the development of new statistical methods to use the data from these arrays for different purposes. The genetic epidemiology of regular voluntary exercise behavior is no exception to this trend. As clear indicators of this trend, the largest reviews to date of twin-, family and gene-finding studies on exercise behavior (Aasdahl et al., 2021; van der Zee & de Geus, 2019; van der Zee, Schutte, et al., 2019) have been published during my PhD project, some of which are included in this thesis. In the same time frame, four of the six genome-wide association studies on physical activity traits (Doherty et al., 2018; Klimentidis et al., 2018; Lin et al., 2018) including sports and exercise behaviors (Doherty et al., 2018; Klimentidis et al., 2018; Lin et al., 2018; van der Zee et al., 2021) saw the light and these have produced the first genome-wide significant genetic variants.

Throughout this dissertation I have employed several methods to further our understanding of determinants of regular voluntary exercise behavior, with an emphasis on genetic factors. In the first three chapters of this thesis, I explore the prevalence and stability of exercise, and review currently available literature concerning genetic epidemiology of exercise. These chapters are followed by an extended-family study, a genome-wide association study (GWAS), and finally a study on the causal effects of personality on exercise, using genetically informed designs. A recurring theme throughout this dissertation was the separation of the total volume of exercise into activities in different domains. The main findings in this thesis are summarized below, followed by a general discussion on the genetic epidemiology of exercise, and my thoughts on future developments in this field.

Prevalence and stability of voluntary exercise behavior

As mentioned in the introduction of this thesis, voluntary exercise behavior is a prime target for behavioral intervention to increase physical activity. Since the health benefits only persist as long as the exercise activities persists, the ultimate goal of such interventions is to develop life-long habits of regular exercise. In other words, the ultimate goal is to not only increase the prevalence, but also increase the *tracking* (i.e. the stability over time) of exercise behavior. The age- and sex-specific prevalence are generally reported in all studies on exercise behavior, however an assessment of tracking is more rare as it requires a large longitudinal dataset. Creating such a dataset is a costly endeavor, especially when the aim is to assess tracking across the lifespan. It is, therefore, not surprising that many previous studies on tracking of exercise focused only on a specific range, such as from childhood to adolescence (Aarnio et al., 2002; Francis et al., 2013; Pahkala et al., 2013; R. Tammelin et al., 2014).

In **Chapter 2** of this thesis, we used a large dataset (N = 43,889) from the Netherlands Twin Register (NTR) to assess prevalence and tracking coefficients across the largest part of the lifespan (ages 8 to 80), with varying distances between the baseline exercise measurement and follow-up (2 to 22 years). We assessed the prevalence and tracking of total weekly volume of exercise as well as of the volume of activities in six specific exercise domains: team-based, solitary, competitive, non-competitive, externally paced, and internally paced. Generalized estimating equations (GEE) were used throughout to correct for family-relatedness in the NTR sample.

The prevalence of exercise increases during childhood and early adolescence, with the highest prevalence found in mid-adolescence, around age 16 for both men and women. The largest contributor to this peak is the prevalence of team-based and competitive exercise activities, which are the dominant exercise

activities in the Netherlands till around age 25. These activities show a strong linear decline after age 16 the slope of which is attenuated around age 40. After age 25, solitary activities, often non-competitive in nature, become the mainstay of the total weekly volume of exercise behavior. The prevalence of these solitary non-competitive exercise activities, remarkably, remains relatively stable from ages 18 to 64. After age 64 the prevalence of non-competitive solitary exercise also starts to decline, though at a slower rate than team-based competitive exercise did during adolescence. Throughout the life span, men have on average 25% higher levels of total exercise than women, but the sex difference is most pronounced in adolescence and early adulthood. This is because the discrepancy in total exercise between men and women during adolescence could be largely attributed to the greater prevalence of team-based and competitive exercise in men.

The tracking of exercise, regardless of type of exercise was moderate to high (2-year tracking coefficients between .38 and .77) with slightly higher tracking coefficients in men. In accordance with previously published tracking studies (Sallis et al., 1996; Telama et al., 2006) we found tracking strength to decrease as the distance between starting age and follow-up age increased. For example, tracking coefficients for total volume of exercise had a median of .57 across all 2-year intervals down to a median of .38 for 22-year intervals. Also in accordance with previous literature we found tracking coefficients generally increased with starting age resulting in lower tracking from childhood into adulthood compared to the higher coefficients within adulthood. For example, in females the 4-year tracking coefficient of the total volume of exercise from age 10 at baseline to age 14 at follow-up was 0.43, whereas the tracking coefficient from age 48 at baseline to age 62 at follow-up was 0.81. However, the increase in tracking strength with increasing age was entirely driven by a sharp increase in tracking for solitary and non-competitive activities. For team-based

and competitive activities, tracking was fairly stable or even gradually decreased across the life span, possibly because the prevalence of such activities decreases. These results suggest that adult exercisers engaged in solitary exercise become increasingly less likely to drop out, and vice versa, that non-exercising adults become increasingly less likely to ever take up exercise, even if it is solitary self-paced and non-competitive.

To conclude, regular voluntary exercise behavior declines with age, but this is largely driven by a decrease in team-based activities. Solitary exercise activities show a much more modest decline in prevalence. The significant female disadvantage in exercise prevalence is moderated by type of exercise and age. Throughout, the moderate to high tracking coefficients confirm that exercise behavior is a rather stable individual characteristic.

Heritability of regular exercise behavior

A large volume of genetic studies has accrued in adult populations that looked at regular voluntary exercise behavior, as well as three other physical activity phenotypes, namely total physical activity (TPA), moderate-to-vigorous physical activity (MVPA), and leisure-time physical activity (LTPA). While TPA, MVPA and LTPA all encompass voluntary exercise behavior, they are much broader in also including non-exercise activities like walking, bicycling or gardening, and activities that are not fully voluntarily controlled by an individual, such as physical activity for transport or manual labor. Total physical activity places no restriction on intensity of the activity, whereas MVPA uses only activities with a minimum intensity (generally activities with an intensity value of more than 3 or 4 times the resting metabolic rate). LTPA again does not select activities based on intensity, but on whether they are performed voluntarily in leisure time and most closely resembles voluntary exercise behavior. The key distinction

being that LTPA also includes all non-exercise or sports performed in leisure time, such as walking or dancing, on top of exercise and sports activities.

Whereas the past studies on exercise genetics have unanimously shown an important genetic contribution to physical activity behaviors, estimates of heritability have varied considerably, even for a similarly defined exercise phenotype (de Geus et al., 2014; Lightfoot et al., 2018). In **Chapter 3** of this thesis, we therefore performed a systematic search and a meta-analysis of adult twin studies on regular voluntary exercise behavior, as well as three other physical activity phenotypes. Twin-studies are particularly well-suited to estimate the heritability of a trait by comparing the difference in within-pair resemblance of monozygotic (MZ) or dizygotic (DZ) twins. A greater resemblance in MZ twins compared to DZ twins suggests additive genetic factors (A) affect the variance in a trait and additional non-additive genetic factors (D), if the difference is more than a factor two. This is because MZ twins share (nearly) 100% of their genetic variants, whilst DZ twins share, on average, about 50% of additive genetic and 25% of non-additive genetic variance. Conversely, if the resemblance in MZ and DZ twins does not differ significantly this suggests that common environmental factors (C) influence the variance of a trait. This is because both MZ twins and DZ twins share all of their common environmental factors, such as parenting styles or prenatal environment, by definition. Finally, the differences between MZ twins can only be ascribed to unique environmental factors (E). Of note, E also includes any phenotypic measurement error.

A set of 27 independent adult twin studies were included, some of which used objective measures (such as accelerometers), others used self-reported surveys. We meta-analyzed the A and C estimates of each study (almost no D was reported) for males and females separately and corrected for the original study sample size.

Contribution of A was significant and substantial in all traits, with the lowest meta-analytic estimates found for MVPA (A_{males} = 0.41, $A_{females}$ = 0.44), and the highest estimates found for TPA (A_{males} = 0.51, $A_{females}$ = 0.48) and exercise behavior (A_{males} = 0.48, $A_{females}$ = 0.51). The contribution of C was only found to be consistently significant in voluntary exercise behavior (meta-analytic estimate 0.10 for both sexes), but not for the broader physical activity traits. No significant sex-differences were found for any of the A and C estimates resulting from the meta-analysis.

In **Chapter 4** of this thesis, we additionally included parent-offspring and sibling correlations from family studies in our review of studies assessing the genetic and environmental contributions to regular voluntary exercise behavior. Twin studies rely on a number of assumptions, one of which is the equal-environment assumption, which states the environmental effects are equal for MZ and DZ twins. Sibling correlations and parent offspring correlations are not relying on this assumption. However, an assumption in estimating A from parent-offspring in the standard family-based design correlations is that the transmission of trait resemblance is entirely due to genetic and not to cultural transmission. In the case of cultural transmission, the behavior of parents directly influences the behavior of the offspring (i.e. parents act as a role-model), thus increasing the influence of common environmental factors. In previous work in the NTR, de Moor et al. (2011) found a minimal (3%) effect of cultural transmission in boys only, despite their participants being adolescents, for whom shared environmental factors are still significant. In keeping, the heritability estimates resulting from the family-studies on exercise that we reviewed in Chapter 4 suggest that those found in twin-only studies are indeed not overly biased.

More alarmingly, however, **Chapter 4** did find spouse correlations to be consistently significant and substantial, ranging

from 0.16 to 0.48. This indicates noticeable assortment for exercise behavior, which violates another assumption of the classical twin design, namely that the trait should not influence an individual's choice of a partner. If such assortative mating is present, and not corrected for, DZ twins in a family may share more than 50% of genetic factors influencing a trait, which in turn will cause an overestimation of the influence of common environmental factors, and an underestimation of additive genetic factors.

In **Chapter 5** we addressed assortment in combination with a final major shortcoming of the twin- and family-studies included in chapters 3 and 4. These previous studies only estimated the effects of A, C and E, but could not simultaneously assess the effects of non-additive genetic factors (D). To achieve this, we used data from 50,690 adolescent and adult participants from 19,543 nuclear pedigrees in the NTR. All family-relations amongst these individuals and within these nuclear pedigrees were known, and this knowledge was leveraged by using the Mendel software package (Lange et al., 2013). Different definitions of 'shared environment' or household effects were tested, namely full household (where C contributes to spousal, parent-offspring, twin- and sibling resemblance), spouse household (where C only contributes to spousal resemblance), sibling household (where C only contributes to resemblance of twins and siblings), and twin household (where C only contributes to resemblance of twins). While still focused on voluntary exercise behavior, we expanded on previous twin-family studies by not only estimating the contribution of the variance components to the total volume of exercise but also to subsets of activities in different domains like team-based or solitary or competitive and non-competitive exercise.

Depending on which definition of household effects was used the broad-sense heritability (h^2, the percentage of variance explained by A and D) ranged from 34% to 41%. Higher heritability estimates were

found for team-based, competitive and externally paced exercise behavior (39% < h^2 < 47%), compared to solitary, non-competitive and internally paced exercise (26% < h^2 < 34%), and were largely attributable to an increased contribution of non-additive genetic factors. Across all exercise domains the contribution of C strongly depended on which definition was tested, with low estimates found under the twin- or sibling household definitions, and higher estimates for the spouse- and full household definitions. For total volume of exercise, for example, the percentage of variance explained by sibling household effects was 4%, by twin household effects 8%, by full household effects 20%, and by spouse household 24%. These large differences mainly reflected the high spousal correlations found across all exercise domains (0.30 < r < 0.50).

Chapter 5 also examined the causes of spousal resemblance for regular exercise behavior, with a particular interest on how they could bias twin models. The only study that had examined spousal resemblance of exercise behavior was published by de Moor et al. (2011), who found that this resemblance was best explained by phenotypic assortment. To re-assess the sources of this spousal resemblance we adopted the approach outlined by van Grootheest et al. (2008) in Chapter 5. Using this model allowed us to test three distinct sources of spousal resemblance in exercise behavior. First phenotypic assortment, which describes preferential mating between partners that more resemble each other's exercise behavior. Second social homogamy, which refers to similar partner selection based on sharing a milieu in which exercise is common. Finally marital interaction, which describes partners growing more similar over time. This could mean that the non-exerciser starts exercising, influenced by the other, or vice versa. Our results suggested marital interaction was a significant contributor in all exercise domain, and there was more evidence in favor of social homogamy over phenotypic assortment. This increases

confidence in twin studies of voluntary exercise behavior that will be most strongly biased by phenotypic assortment, to a lesser extent by social homogamy, and even less by marital interaction.

To conclude, results of chapters 3-5 are in line with those of previous twin and family studies regarding the heritable nature of regular voluntary exercise behavior as well as broader physical activity traits in adulthood, and the near absence of lingering effects of being raised in a shared family environment. Like so many other behavioral traits, "about half of the variance" in adult exercise behavior can be explained by heredity. This heritability, however, is a result of both additive, and non-additive effects, particularly but not exclusively for team-based, competitive and externally paced exercise. Marital interaction seems to be consistent source of influence on spousal resemblance in regular voluntary exercise behavior and there is likely an additional role for social homogamy. These effects are of greater influence than phenotypic assortment.

Gene-finding in regular voluntary exercise behavior

As mentioned in the introduction of this thesis discovering the genetic variants underlying the heritability of regular voluntary exercise behavior has proven to be difficult, despite the well-established moderate-to-high heritability. A large number of candidate-gene studies have been performed in which a single or a few genetic variants were assessed for an association with regular exercise or broader physical activity traits (see review in Chapter 6 but also Aasdahl et al. (2021)). Only very few candidate genes, however, have been independently replicated and then not even always at the level of a specific variant but rather at the gene level. A genome-wide association study (GWAS), where millions of genetic variants are tested simultaneously, is a more robust and hypothesis-free method for identifying genetic variants. The first three GWASs on regular

voluntary exercise behavior, published by de Moor et al. (2009), Kim et al. (2014) and Lin et al. (2018), did not yield any significant results. The two GWASs that followed, published by Klimentidis et al. (2018) and Hara et al. (2018), did yield the first genome-wide significant hits for sports and exercise behaviors. Genetic variants for specific types of exercise activities, however, were not tested.

In **Chapter 6** of this thesis, we performed a GWAS of the total weekly volume of exercise, as well as a GWAS for activities in the team-based, solitary, competitive, non-competitive, and externally/internally paced domains as defined previously. We used a sample of the NTR (N = 14,626) for which genome-wide data in the form of millions of single nucleotide polymorphisms (SNPs) was available as well as survey data on exercise activities.

No genome-wide significant SNPs were found for total, team-based, solitary, competitive, externally paced and internally paced exercise. For non-competitive exercise five novel genome-wide significant SNPs were found. Four of these variants (*rs11817437*, *rs1276262*, *rs11812727*, *rs34897771*) were located in an intergenic region on chromosome 10 and have not been found in any previous genetic study of exercise behavior. One of these significant variants (*rs147244851*) was located in the *Glypican-5* (*GPC5*) gene. Though this gene has previously been found to be associated with sedentary behavior (Comuzzie et al., 2012), the variant found in our GWAS on exercise behavior was different from the variant associated with sedentary behavior. Since there is no evidence for either our exercise variant, or the previous sedentary behavior variant affecting gene expression, we cannot say with certainty that the direction of the effect found in these two GWASs is congruent.

We used our suggestive ($p < 5\times10^{-6}$) GWAS results to perform several functional annotation analyses. Through these analyses one of the variants (*rs62620995*) found for total volume of exercise was

identified as a missense (gene-product altering) variant in the *DCSTAMP* gene. This gene plays a crucial role in bone-development (Kukita et al., 2004), furthermore the same variant found in our GWAS has been shown to affect *DCSTAMP* gene expression and osteoclast (bone-resorbing cells) morphology (Laurier et al., 2017). Across the different types of exercise activities, the suggestive GWAS results were consistently enriched for genes that were related to a number of red blood cell phenotypes, such as glycated hemoglobin levels, mean corpuscular volume, and heme metabolism.

In chapter 6 we also aimed to replicate the 103 genetic variants identified in previous candidate-gene studies and test all these variants (or a genetic proxy of these variants) for association. This replication used a much more relaxed *p*-value than the GWAS, namely 0.05/103, i.e. the 'standard' α corrected for multiple testing. Strikingly, no significant evidence was found for replication of any of the previously reported genetic variants.

To conclude, in chapter 6 we identified novel genetic variants for regular voluntary non-competitive exercise activities specifically. Results of this chapter tentatively suggest bone development and red blood cell physiology to be biological determinants of voluntary exercise behaviors.

The pathways from genetic variation to regular voluntary exercise behavior

The recent discovery of significant genetic variants for physical activity traits, including regular voluntary exercise behavior, allows us to use new methods that leverage these variants to test causal hypotheses on the role of potential 'intermediate phenotypes' on the path from genes to exercise behavior. Personality, a trait known to be heritable, is one of these potential determinants of regular voluntary exercise. There is robust evidence for an association between these

two behavioral traits (Rhodes & Smith, 2006; Wilson & Dishman, 2015). The oldest, and still predominant perspective on this association is that it reflects a causal effect of personality traits on exercise behavior (Flemming, 1934), though an inverse causal effect has also been hypothesized (Stephan et al., 2014), and there is much room for genetic and environmental confounding. Testing the causal hypothesis, however, is not possible with traditional methods such as a controlled trial, as we cannot directly experimentally modify a subject's personality. This is where new genetically informed designs come in, including those based on twin-families and/or GWAS derived variants.

In **Chapter 7** we used a number of these designs to test the causal effects of the 'big five' personality traits (neuroticism, extraversion, agreeableness, openness and conscientiousness) on the total volume of exercise behavior, as well as on the volume of competitive/team and non-competitive/solitary exercise activities. First, we assessed the cross-sectional and longitudinal associations between the five personality traits and the various exercise behaviors. Second, as a direct test of causality we regressed the intrapair differences in MZ twins in exercise behaviors on the intrapair difference in the personality traits, thereby controlling for possible genetic and shared environmental confounding. Third, we tested a crucial requirement to uphold the causal hypothesis, namely that all genetic and environmental correlations between personality and exercise behavior would be significant. This was done in two different twin-sibling models, each with its own strengths. In the first model, to reduce impact of measurement error, we used the longitudinal data (two observations per participants) in a bivariate model as multiple indicators of the 'true phenotype'. In the second model, to reduce impact of reverse causality, we used longitudinal data in a prospective design, such that personality at baseline was a predictor for exercise behavior at follow-up. We separately re-assessed the genetic

correlation in independent data, using the available leave-NTR-out GWAS summary statistics for personality (de Moor et al., 2015; Lo et al., 2017) and for sports and exercise behavior (Klimentidis et al., 2018). Finally, the summary statistics from these GWASs were used to generate polygenetic scores (PRS) for personality traits and exercise behavior to be used as genetic instruments in a direct test for causality using the Mendelian randomization direction of causation (MR-DoC) model (Minică et al., 2018).

Results provided the strongest support for the causal hypothesis, i.e. by resisting falsification, for the association between extraversion and total volume of exercise, closely followed by the association between conscientiousness and non-competitive/solitary exercise and conscientiousness and total volume of exercise. Of note, we failed to show a causal effect of neuroticism on exercise behavior, which is in line with a previous cross-lagged panel analysis by de Moor and de Geus (de Moor & de Geus, 2018). Interestingly there was some evidence for a positive association between openness and non-competitive/solitary exercise, but no association between openness and total volume of exercise, likely due to the (non-significant) negative associations between openness and team-based exercise. Reverse causal effects of exercise behavior on extraversion, conscientiousness and openness were found consistently, although the effect sizes for such reverse causality were consistently smaller in our strongest estimator of the causal effect, the MR-DoC model.

From the results of chapter 7 we conclude that there is likely a bidirectional causal relationship between the personality traits extraversion, conscientiousness and openness to experience and regular voluntary exercise activities.

Discussion

The main conclusions and contributions of this thesis can be divided into three main topics, discussed separately below. First, the prevalence and stability of exercise, which mainly concerns chapter two, and the other chapters to a lesser extent. Second, the genetic epidemiology of voluntary exercise, where chapters 3, 4 and 5 mainly address the heritability, chapter 6 addresses gene-finding, and chapter 7 serves as an example of how this knowledge can be used to further study the pathways from genes to voluntary exercise behavior. Finally, the added value of a subdivision of total weekly exercise into activities in specific exercise domains, which has been a common theme throughout this thesis.

Exercise prevalence and stability

The results of chapter 2, as well as exercise descriptive statistics in other chapters provide yet more evidence that the weekly volume of regular exercise decreases with age and depends on sex. The peak of exercise behavior is in mid-adolescence, after which it decreases sharply until at about age 40 where it shows a more gradually slowing downward trend that re-accelerates around age 65. Males tend to exercise more than females at all ages although the rate of decrease after mid-adolescence is greater in males. These findings are by no means novel, and this pattern is a well-known characteristic of exercise behavior (Azevedo et al., 2007; Eime et al., 2016; Hirvensalo & Lintunen, 2011; R. Tammelin et al., 2014). The results of this thesis do, however, shed some new light on this pattern by splitting exercise activities into the various domains introduced previously. Our results indicate that the strong decrease in exercise after mid-adolescence is almost entirely due to a decrease in team-based competitive exercise in both males and females, whereas solitary non-competitive exercise remains relatively stable. This is in line with what was stated earlier by

Telama et al. (2005) based on work by Burton (1988): "there are large numbers of young athletes who would prefer to continue competitive sport but do not do so because they are screened out because of low achievement or because they perceive a sport as too demanding in the relation to their perceived ability". The strong emphasis on the *perceived* demand and *perceived* ability is further in line with a pathway model by de Geus (2020), which suggests a feedback loop between exercise behavior and perceived ability, where high perceived ability acts as positive reinforcement for exercise behavior, but low ability as punishment.

The above patterns in exercise prevalence suggest that the transition from competitive team-based sports to solitary exercise is currently a 'weak link' from a public health perspective. This is where a major drop-out occurs. Social roles may start to interfere with team-obligations and/or the level of organized team sports for adults is not as advanced as that for children and adolescents. It is not clear why many exercisers do successfully switch to other exercise activities, whereas others don't. Interestingly, the perceived value of competitive elements does not seem to change. We found a decreasing correlation between team-based and competitive exercise in parallel to an increasing correlation between solitary and competitive exercise throughout adulthood. Individuals favoring a competitive element tend to switch from a team-based (e.g. field hockey or soccer) to a competitive solitary exercise (e.g. tennis or squash) where they still obtain the perceived benefits of competition.

Exercise behavior that survives the transition into adulthood shows an increasingly higher temporal stability. Adult regular exercisers seem to have found a set of exercise activities that is sufficiently reinforcing to be maintained, possibly because it optimally matches their genotypes.

Genetic epidemiology of exercise

Previous findings suggest that the heritability of exercise behavior increases from childhood to adolescence, with a peak at around 18 years (Huppertz et al., 2016; Vink et al., 2011). As reviewed in Schutte et al. (2018), this occurs in parallel to a large change in the contribution of shared environmental influences. During childhood, sibling resemblance in exercise behaviors is strongly determined by the shared environment in which they grow up, but this shared environmental influence strongly wanes during adolescence. The strong decline of contribution of shared environmental factors may be due to the parental motivation increasing the volume of exercise behavior in their children (Beets et al., 2010). The effect of this parental motivation does decrease over time, giving way to increased influence from peers and coaches (Chan et al., 2012).

When the adolescents become young adults, heritability appears to decrease again to reach the levels reported in chapters 3 and 4, where we conclude that across the full adult life span about half of the variance in exercise behavior is contributable to genetic factors. Fine-grained analysis of the adult trajectory (as is available for children and adolescents) is currently lacking. However, various indicators suggest that this reflects a decrease in heritability from the peak in adolescence to lower levels in later stages of adulthood where the decrease becomes asymptotic. For example, the estimate of 48% from the meta-analyses in chapter 3 is higher than our finding in chapter 5, where we find a slightly lower heritability of roughly 40%, depending on the type of exercise. This is likely due to the large proportion of studies in the meta-analysis, with a relatively low mean age (~28) compared to the extended twin-family study in chapter 5 that had a mean age of 40 years.

This decreased influence of genetics is paired to a gradual increased influence of person-specific environmental influences.

There is a huge number of factors that can come into play in exercise behavior in adulthood such as entering a stable relationship, having children, job demands, caretaking obligations, or health problems (Dishman et al., 1985; Seefeldt et al., 2002; Simonen, Videman, et al., 2003). Furthermore, as identified in our extended twin-family study, we find a significant contribution of spouses on each other's exercise behaviors. Assuming individuals do not share a household with their spouse before adulthood and given the significance of marital interaction we find in chapter 5, the spousal influence on exercise behavior after partnering up may slowly increase up to 24% of explained variance.

Combining all results obtained in the NTR cohort I would like to propose an extension to the figure by Huppertz et al. (2012) showing the change in relative contributions of genetic and environmental factors to the total volume of regular voluntary exercise behavior as a function of age.

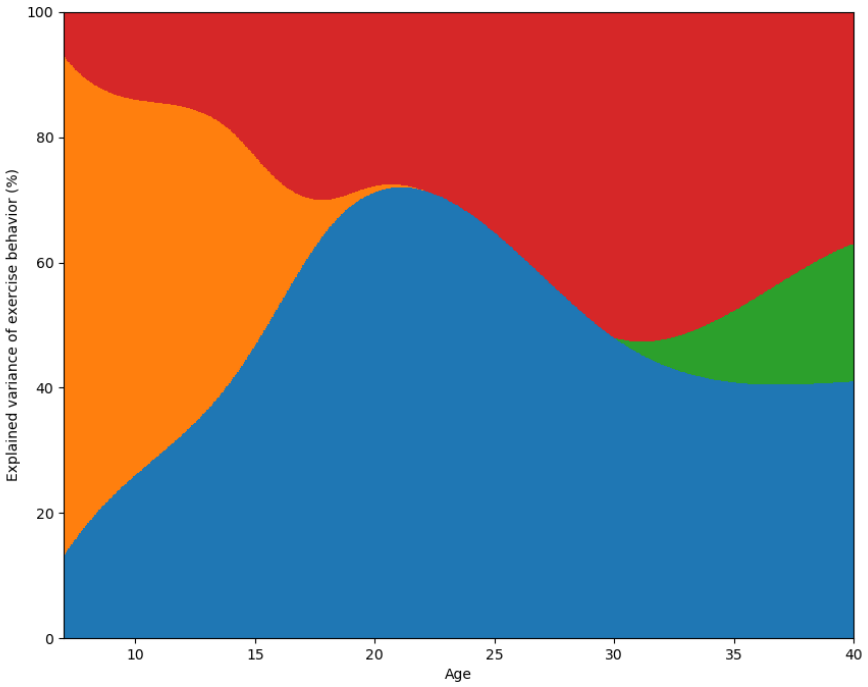

Figure 8.1. Percentage of variance in regular exercise behavior explained by unique environmental factors, environmental factors shared by siblings, genetic factors, and environmental factors shared by spouses over time.

It is important to note that the genetic factors displayed in my Figure 8.1 includes both additive, and non-additive genetic factors. Due to the limitations of the classical twin model, only one of common environmental factors, or non-additive genetic factors can be included in the model. If non-additive genetic factors are not specifically defined in the model, their contribution to the variance will end up in the additive genetic component. Given our finding that about half of the heritability (and even more so in team-based and competitive exercise) is attributable to non-additive genetic factors it seems likely that these factors also play a role earlier in life. More research is needed to accurately separate non-additive from additive genetic factors

throughout the lifespan, so for now the blue-colored area reflects broad sense heritability.

Four main tenets are visualized in this figure. First, the earlier mentioned strong decrease in the contribution of environmental factors shared by siblings from childhood (age 7) to late adolescence, to the point where these factors are no longer a significant contributor (around 16 years). Second, a strong increase in contribution of genetic factors up to a peak heritability of 70% in late adolescence and early adulthood (around 20 years) that levels off to 40% at age 40. Third, the near linear increase in the influence of unique environmental factors across the life span. Fourth, spousal factors become an important component of these environmental factors from age 30 onwards and form a substantial (de)motivator to engage in exercise after age 35.

There are different possible explanations for the sharp increase in the relative importance of genetic factors during adolescence followed by a gradual decrease to middle adulthood. First, the changes in the percentage of variance explained by genetics may simply be a result of changes in the importance of common and unique environmental factors. In other words, the *true* genetic effects may not change, only their relative contribution. However, a true increase in absolute genetic variance was demonstrated from childhood to adolescence (Huppertz et al., 2016) making this explanation unlikely. That unique environmental influence start to have a cumulative effect from adolescence towards middle adulthood makes more sense and a change in the relative contribution of E and A could explain part of the gradual decrease in heritability in this period.

However, the peak in adolescence and the decline in adulthood could also be related to the change in the prevalence of team-based and competitive exercise activities. It is very high in adolescence but wanes in adulthood. Because we have found team-based and competitive exercise to have an overall higher heritability in chapter 5,

decreased participation in primarily these types of exercise activities would lead to a proportional decrease in heritability of total volume of exercise. Second, the decreasing heritability may be a direct result of different genetic effects coming into play at different ages. In other words, a *true* change in the genetic effects may occur. Such a change may be related to a behavioristic model for the genetic determinants of exercise behavior presented in de Geus (2021). Core determinants of exercise behavior are grouped in two instrumental conditioning loops continuously influencing the adoption and maintenance of regular exercise activities. One loop is related to the affective responses during and after exercise, which is linked to genetic effects on the neurobiological impact of exercise on reward/punishment circuits; the other loop is related to the rewarding effects of being able to perform well on a valued activity, which is linked to genetic effects on core elements of exercise ability including fitness, strength, low injury sensitivity and body composition.

The difference in heritability estimates found for different types of exercise activities may be the result of different genetic effects on these types of exercise. Team-based competitive activities would be favored by genetic variants influencing exercise ability whereas solitary and non-competitive activities would be favored by genetic variants influencing the affective response to exercise. In other words difference in heritability estimates can reflect that these activities are, at least partly, influenced by different genetic variants. If this is the case, then we would expect (1) to find different GWAS results for team-based and competitive exercise compared to solitary and non-competitive exercise, and given the changing proportions of different types of exercise behavior across various ages (2) to find different GWAS results for total exercise behavior depending on the age of the sample.

In line with these expectations, we find different results for the different types of exercise activities in our GWAS in chapter 6, especially when comparing team-based competitive exercise activities to solitary non-competitive activities. Furthermore, we find no direct overlap between our GWAS results in a sample with a mean age of 31 and the self-reported total volume of exercise behavior in the largest GWAS available for self-reported sports and exercise behavior in the UKB, which is a sample aged 40-69 years. Nor did we find overlap with the variants detected in the second largest GWAS by Hara et al. (2018) in participants with a mean age of 55. Furthermore, the genome-wide significant variants found in the UKB by Klimentidis et al. (2018) appeared to be largely brain-related. In contrast our GWAs in younger population, found variants related to red blood cell phenotypes.

It is attractive to think that our GWAS tapped more into the genetic variants influencing ability whereas the UKB GWAS tapped more into genetic variants influencing the affective response to exercise. There is a clear association between red blood cell physiology and exercise ability (Mairbäurl, 2013) mainly due to their role in oxygen-delivery to peripheral muscle. A case in point is the story of Eero Mantyranta, who had stellar biathlon capability due to a rare mutation in the EPO receptor gene (Enriquez & Gullans, 2012). It is therefore not surprising that red blood cell increasing substances such as erythropoietin (EPO) are banned from professional sports as performance-enhancing drugs (World Anti-Doping Agency, 2021). Thus it seems very plausible that individuals with a genetic makeup for an increased number (or improved physiology) of red blood cells perform better compared to their peers. The perception of higher ability would in turn make these individuals more likely to continue exercising as part of (competitive) team sports activities, in line with de Geus (2020) as discussed earlier.

Clearly, this is utter speculation. There are multiple other possible sources of the differences between the handful of exercise GWASs, most notably the modest sample size in our NTR GWAS. Apart from insufficient power, there is potential bias due to selective participation (both in the UKB and NTR cohorts) resulting in population stratification and poor harmonization of instruments used to establish voluntary exercise participation. The different findings might also be the result of different moderation of genetic variants across different countries (i.e., gene-environment interaction). Larger GWAS on the various types of exercise in different populations and at different ages are very much needed to progress our understanding of the molecular genetic basis of voluntary engagement in regular exercise behaviors and its change across the lifespan.

Exercise domains

The main exercise phenotype used throughout this thesis was the total amount of sports and exercise activities weekly, recast as a measure of energy expenditure on these activities. This involves the use of fixed intensity codes for specific activities based on the average values obtained in past experimental studies (Ainsworth et al., 2011; Ridley et al., 2008). This had the advantage of making our exercise phenotype comparable to those used in a large number of studies in the extant literature. A disadvantage of this measure, compared to e.g. number of minutes weekly spent on exercise, is that it introduces measurement error because individuals may perform these activities at different actual intensity levels based on differences in motivation or efficiency related to body composition and/or motor competence. However, acknowledging this disadvantage, we could now at least compare our results to a large body of past work that also used a METminutes-based score.

The additional use of energy expended on team-based, solitary, competitive, non-competitive, externally paced and internally paced exercise activities was more challenging in this regard. The inclusion of these separate exercise domains was inspired by van der Mee et al (2017) and has become a common theme throughout this thesis. As a consequence, many results presented in this thesis, i.e. those related to these domains specifically, cannot be directly compared to existing literature, as to my knowledge almost no previous genetic studies have used these domains. However, I believe the results presented in this thesis strongly justify the use of these separate domains, as they yield differential and novel findings, compared to total volume of exercise alone. Most strikingly was the differential course across the life span of team-based, competitive and solitary, non-competitive activities in chapter 2 as well the larger broad sense heritability of the former in chapter 5 due to a more substantial non-additive component.

Do all these domains need to be included in the analyses? It may depend on the application. Because of the clear meaning the different domains have for professionals and the clear role these may play in possible interventions it is my recommendation to adhere to these domains when the aim of future research is directly aimed at developing or improving exercise interventions. For genetic research, we did note substantial redundancy between team-based, competitive and externally paced activities on the one hand, and solitary, non-competitive, and internally paced activities on the other hand. Internally (self) paced versus externally paced, in particular, did not add a lot to the domain solitary vs team-based. However, when examining the role of the genetics of executive functioning on exercise preference this distinction did prove to be meaningful (van der Mee et al., 2017). Somewhat more distinction exists between the team-based vs solitary domain and the competitive and non-competitive domain, although substantial overlap was found for these dimensions too. Even

so, the inclusion of the full set of team-based versus solitary and competitive versus non-competitive may be maintained when research concerns, for instance, the role of personality in exercise (H. J. Eysenck et al., 1982).

Whereas some research questions may require all three exercise domains to be used in full, evidence from both the phenotypic and genotypic analyses performed throughout this thesis suggest that two underlying factors capture the six METminutes values across the three exercise domains rather well. These latent factors can be approximated by using the team-based versus solitary activities but an actual factor score computed across the six different exercise phenotypes, as used in chapter 7 may be optimal. On top of the inclusion of the separate domains or the two factor scores, the total volume of exercise behavior should also be included since (1) the correlations between total volume and the two factors or the full set of the six METminutes scores are only moderate and (2) the vast majority of previously published literature only includes total volume of exercise behavior, as discussed earlier, thus allowing for a more direct comparison to past findings.

Future directions for exercise phenotyping

Throughout this thesis self-report, or parent-report in children included in chapter 2, was used to measure the time spent on all exercise activities. As discussed in multiple chapters of this thesis, voluntary exercise activities are salient and thus reliably recalled by self-report in surveys. Total physical activity on the other hand is much harder to accurately recall retroactively, it is too easy to significantly over- or underestimate the time spent being physically active in a day, let alone a week. It is therefore not surprising that the accelerometer (and pedometer to a lesser extent) has all but replaced the self-report survey in studies on physical activity (Silfee et al., 2018).

Nowadays the vast majority of the population carries highly valid accelerometers with them everywhere all the time, at least one in (nearly) every smartphone, and sometimes additional ones in devices such as fitness bands or smartwatches. The large amount of data generated by these devices provides a lot of opportunity for future research. Major challenges, however, are *to obtain* the data from the often commercial suppliers like Fitbit, Google, Apple or Polar, and *to standardize* across the many proprietary versions of the hard- and software that regularly spring into life. For this reasons, many researchers have resorted to dedicated research devices for accelerometer signals, of which the Actigraph seems most popular, paired to scoring standards and open software solutions in R to do so (Migueles et al., 2021). Unfortunately, where total physical activity was harder to measure through self-report, voluntary exercise behavior cannot be assessed by relying solely on an accelerometer. Since the accelerometer only records raw movement on three axes, it is perfectly suited to measure total physical activity, however without a record or log about what type of activity was performed (and why), there is no easy way to tell which bout of activity is attributable to voluntary regular exercise behavior.

The future is bright here though, as innovative machine learning methods become increasingly capable at recognizing patterns in vast quantities of unstructured data, making them very promising methods to accurately detect exercise bouts from raw accelerometer data (Hoogendoorn & Funk, 2018). Ongoing research to make these classifiers more and more accurate are making it feasible to use accelerometers in large-scale exercise studies in the future. The benefits of this are plenty, for example we can more accurately estimate the exact amount of time an individual is engaged in exercise, and how often. Additionally, as these models develop, they will inevitably produce a better estimate of the true energy expenditure

during exercise compared to survey-based data in conjunction with MET-constants (Freedson et al., 2011; Lawal & Bano, 2020; Staudenmayer et al., 2009; Yen & Lin, 2020).

Another exciting avenue for research is brought about when accelerometers are combined with short self-report queries that are sent out on an hourly basis as part of people's daily lives, also known as ecological momentary assessment (EMA). Smartphone-based EMA can measure an individual's intent to engage in, attitudes towards *voluntary* exercise, and experienced enjoyment. Combining EMA with accelerometers can accurately assess the differences in people's reported or intended exercise activities, acute psychological effects of exercise, and momentary determinants of exercise, together with the objective volume of exercise they engage in (Dunton et al., 2015; Liao et al., 2017; Liao et al., 2015; Maher et al., 2017).

In spite of these exiting developments, survey-based research into exercise behavior will retain its value in the foreseeable future. No matter how good the accelerometers and the accompanying (machine learning) models become at estimating time and energy spent on exercise, the scale of the samples required for genetic analyses may not be easy to obtain with these methods due to the challenges mentioned previously. Also, various trait-like characteristics are still well captured by surveys like personality or attitudes. One other relevant trait that we recently tackled is an individual's exercise *liking*. In over 157,000 individuals from the UK Biobank, we sought to complement and extend previous findings on the genetics of physical activity behaviors by performing genome-wide association studies of liking of several exercise-related behaviors plus an additional derived overall activity-liking trait. We identified a total of 17 loci, along with an additional eight for the overall trait, only some of which overlap with loci previously identified for physical activity behavior. Replication in over 7,000 adults from the NTR showed directional consistency in 13

out of 17 loci (Klimentidis et al., 2021). If the aim of a future study is to aid the development of more personalized exercise interventions, an individual's exercise liking and attitudes towards the behavior are, in my opinion, at least as important (if not more so) as the objective energy expenditure.

Future directions for finding molecular determinants of exercise behavior

As sample sizes increase, future GWASs will be able to tell us more about the complex genetic underpinnings of exercise behavior, particularly when cohorts combine forces in meta-analyses. Throughout this thesis I have shown that these meta-analyses would do well to take the changing composition of total exercise behavior into account. 'Total exercise' at age 20 is not strictly the same behavior as 'total exercise' at age 60, which may make meta-analysis of studies across different ages more challenging. Additionally, as discussed previously, a GWAS of exercise activities is likely sensitive to selection bias compounded by population stratification. Luckily the development of genetic methods does not seem to be slowing down and new methods to tackle issues like this are being developed or have already been developed. For example to tackle the issue of different compositions of total volume of exercise across the various dimensions, future research could employ the multivariate GWAS meta-analysis approach as described by Baselmans et al. (2019). Additionally, to tackle the issue of stratification or indirect genetic effects future research could use a within-sibling GWAS (Howe et al., 2021).

Variations in the DNA code, the focus in this thesis, are not the only possible molecular underpinnings of voluntary exercise. There is likely a strong relationship between epigenetic mechanisms and voluntary exercise behavior as well (van Roekel et al., 2019; Voisin et al.,

2015). Epigenetics does bring up the thorny issue of causality. In genetics, the direction of causality is simple, as genetic variants can (indirectly) affect behavior in many ways but the reverse is near impossible (barring extremely rare genetic mutations due to things like radiation). In epigenetics, this is not as simple because epigenetic changes such as DNA methylation or histone modification may well have a causal effect on exercise behavior, but the reverse may just as well hold, or the effect may be bidirectional creating an epigenetic feedback loop (Lightfoot, 2020). Further complicating the epigenetics of exercise is the fact that epigenetic profiles are different in different cell types (Maurano et al., 2012). Nevertheless, I believe epigenetics has great potential in exercise behavior as it can lead to more insights into the molecular underpinnings of the health benefits of exercise. Longitudinal epigenetic assessments combined with longitudinal measures of (objective) exercise behaviors can likely aid in distinguishing these two directions of causality. Using twin samples this can be further aided by an MR-DoC design.

Another avenue through which we may well learn more about the biological underpinnings of the health benefits of exercise behavior is the gut-microbiome. The interest in the complex ecosystem of bacteria and other micro-organisms in the gut is relatively young and the immense diversity of the gut-microbiome with over 3 million unique genes from over 1000 unique bacterial species (Qin et al., 2010) poses significant challenges for future research. This diversity is in turn affected by a combination of host genetics (Kurilshikov et al., 2021), and numerous environmental factors (Rothschild et al., 2018), and these effects can vary between species, posing further challenges. Despite these challenges, however, numerous studies in both humans and animals have already found evidence for associations between exercise and the gut-microbiome (see Mailing et al. (2019) and Dalton et al. (2019) for review).

To summarize, I think there is a lot to look forward to in the field of exercise genetics. With the combined growth in popularity of daily-wear accelerometers, advancements in methods to accurately label this data and advancements in EMA methods, there is a wealth of new data just on the horizon. By combining this data with novel and increasingly bias-resistant genetic methods in ever increasing sample sizes, and expanding from the genomic to other multi-omics levels, we stand to learn a lot more about the molecular determinants of exercise behaviors. Combining these results across cohorts in large-scale meta-analyses will yield the required statistical power to detect the small effects that are inherent to any highly polygenetic behavior. If, on top of that, future research also analyzes the different types of exercise activities separately, this will likely not only yield more scientific findings, but also provide avenues for personalized interventions that have higher success rates than those currently on offer.