Supplementary Methods S1

Sex differences in genetic architecture of complex phenotypes

Methods section

1. Lifestyle

-*Ever smoked:* Both ANTR and YNTR subjects were classified as never smokers when they reported they never smoked or when they tried smoking a few times but never smoked regularly [1].

-*Current smoking:* The answer categories to the question "How often do you smoke now?" were collapsed into current smoking yes (I smoke once a week or less, I smoke a few times a week, not every day, I smoke once or several times a day) and current smoking no (I've never been a regular smoker, I've quit smoking) both for ANTR and YNTR cohorts [2].

-*Nicotine dependence:* The Fagerstrom Test for Nicotine Dependence consist of 6 items and produces a score ranging from 0 to 10 with higher scores indicating more nicotine dependence [3-4]. The FTND score (among ANTR smokers and ex-smokers) was collapsed into a dichotomous variable (FTND >= 4 versus FTND < 4).

- *Cannabis use:* In the ANTR, subjects were asked at what age they initiated cannabis use and the answers were re-coded in the variable 'cannabis initiation', with two possible categories; 'yes' when a subject initiated cannabis use (irrespective of age), or 'no' when a subjected never initiated cannabis use [5]. In the YNTR the Dutch Health Behavior Questionnaire (DHBQ) was included. Subjects were asked whether they ever used cannabis and, if so, how many times. Responses were re-coded into the variable 'initiated cannabis' with two possible categories: 0, when a subject never used cannabis; and 1, when a subject had ever used cannabis [6].

-*Ever alcohol* in YNTR: Have you ever used alcohol was re-coded to never (never and 1or 2 times) versus ever (3 times or more).

-*Regular drinking* in ANTR: Regular drinking defined as: >= 21 glasses per week for men and >= 15 glasses per week for women (conform Nationaal Kompas Volksgezondheid www.nationaalkompas.nl (dutch)).

-*Weekly drinking* YNTR: “How often do you drink an alcohol beverage”: answer categories were collapsed into: once a week or more often (answer 5,6,7) versus a few times per month or less (answer 0,1,2,3,4).

-*Alcohol problems* in ANTR : Alcohol problems were assessed by CAGE. The CAGE, originally developed as a screening instrument in medical settings [7], is a short questionnaire to obtain an indication of the presence of Symptoms of AAD [8-9]. Its four items are: Have you ever felt that you should Cut down your drinking?; Have people Annoyed you by criticizing your drinking?; Have you ever felt bad or Guilty about your drinking?; Have you ever had a drink in the morning (as an ‘Eye opener’) to steady your nerves or get rid of a hangover? The items can be answered with ‘yes’ or ‘no’. In 2009/2010, there were two yes categories: ‘yes, during the last year’; ‘yes, not during the last year’ that, for these analyses, were re-coded into one, since its combined frequency was comparable to the frequency of the single yes categories in earlier years. If one of the four CAGE items was missing (N = 296), the missing value was imputed based on the mean of the three other answers. For individuals with two or more missing items (N = 187) and those who reported to never have drunk any alcohol (N=332), the CAGE was set to missing. For the current study, we analyzed the number of ‘yes’ items as 0, 1, 2 and 3+, combining the 3 and 4 yes positive answers as as the number of individuals who answered four items positively, was low (N=56).

-*Early alcohol initiation at age 13-15* (in YNTR sample) was defined as ever having used alcohol, to which response categories were ‘no’, ‘a few times’ and ‘yes’. The categories ‘a few times’ and ‘yes’ were collapsed, resulting in a binary variable [10].

-*Coffee:* The ANTR survey of 2000 contained the question: ‘On average, how many cups of caffeinated coffee do you drink in one day?’ A large part of the subjects reported drinking no coffee at all causing a skewed distribution. Therefore, the data were divided in categories: zero to two cups per day, three to five cups per day and six or more cups per day [11].

-*Exercise participation* in ANTR was measured with a number of questions. The first question was 'Do you participate in exercise regularly?'. This question could be answered with Yes or No. If the participants responded affirmative, further information on type, frequency and duration of exercise was gathered. All exercise activities were assigned a metabolic equivalent value, using Ainsworth's Compendium of physical activity [12]. A metabolic equivalent score of 1 corresponds to the rate of energy expenditure when at rest (1 kcal/kg/h).
Exercise participation was defined as a dichotomous variable, classifying participants as either regular exerciser or non-exerciser [13]. A cutoff criterion of exercising at four metabolic equivalents or more for at least 60 minutes a week in the recent year was used to classify participants as regular exercisers. Exercise participation was assessed in ANTR surveys 1991, 1993, 1995, 1997, 2000, 2002 and 2004. For each twin pair, valid exercise data from the most recent survey were selected.

2. Emotional and Behavioral Problems

Obtained from longitudinal surveys in adults (ANTR, 18-65 years)

-*NEO personality scales:* Personality scores for the five factors Neuroticism, Extraversion, Openness to Experience, Agreeableness and Conscientiousness were assessed with the 60 items of the NEO-FFI (12 items per factor) [14-15]. Items were answered on a 5-point Likert-type scale ranging from strongly disagree (1) to strongly agree (5). Summed scores were computed for all five personality dimensions (after reversing negatively keyed items). If more than 10 items were missing, summed scores were not computed. If less than 10 items were missing, the missing items were given the neutral option (2). Sumscores for each dimension could range from 12 to 60. Personality data came from ANTR survey 7 (2004).

-*Fear* in ANTR: Blood-injury fear (e.g., avoidance of hospitals) and Social fear (e.g., avoiding talking to people in authority) were measured by the Dutch version of the Fear Questionnaire (FQ) [16]. Both domains are represented with 5 items. Subjects were instructed to indicate on a nine point scale (0 'would not avoid it' to 8 'always avoid it') how much they would avoid certain situations because of fear or other unpleasant feelings. Scores were summed across subscales. A score was only calculated if at least 4 out of 5 items of a subscale were answered. Missing answers and double entries were substituted by the mean item score. The sum scores were transformed to a dichotomous measure.

-*Obsessive compulsive symptoms* (OCS): OC symptoms were measured by 12 items of the Padua Inventory [17], translated into Dutch, revised and validated by Van Oppen et al. [18]. Items were chosen from each OC subscale of the Padua Inventory Revised. The sensitivity and specificity for the 12 items to detect OCD were 0.74 and 0.72 respectively, when comparing a group of OCD patients with clinical controls [19].

- *Borderline personality features* were measured by the Dutch translation of the Personality Assessment Inventory-Borderline Features scale (PAI-BOR) [20-21]. PAI-BOR items tap features of severe personality pathology that are clinically associated with borderline personality disorder and consists of 24 items that are rated on a four-point scale (0 to 3; false, slightly true, mainly true, very true). The PAI-BOR was scored according to Morey's test manual [20], which states that at least 80% of the items must be answered to calculate a sum score and that missing and ambiguous answers should be substituted by a zero score.

-*Anxious depression* –Anxious depression was measured with the Young Adult Self Report (YASR) [22-23].

-*State-Trait Anxiety:* Anxiety was measured with the Dutch translation of the Spielberger State Trait Anxiety Inventory – Trait version (STAI) [24-25].

- *State-Trait Anger*: Trait anger was measured with the Dutch adaptation of the State Trait Anger Scale (STAS) [26-27]. The scale is designed to assess the frequency of which an individual experiences the state anger over time and in response to a variety of situations. The 10-item STAS-adaptation was scored on a 4-point likert scale (1-4; almost never, sometimes, often, almost always). Participants were asked to indicate the extent to which an item occurred in their everyday lives. The items were scored according to the test manual, which states that at least 80% of the items must be answered to calculate a sum score and that missing values or ambiguous answers should be substituted by a score of 2 (sometimes true).

-*Loneliness* was assessed with the short scale for measuring loneliness in large surveys, developed by Hughes et al.[28]. The three items that compose this scale were selected from the R-UCLA Loneliness Scale [29]: ‘‘How often do you feel left out’’, ‘‘How often do you feel isolated from others’’, and ‘‘How often do you feel that you lack companionship’’. Response categories were: ‘‘Hardly ever’’ (1), ‘‘Some of the time’’ (2), and ‘‘Often’’(3). The data were analysed as an ordinal measure with 4 categories.

-*Sensation seeking* was measured with the Dutch translation of the Sensation Seeking Scale (SSS) from Zuckerman [30-32]. The sensation seeking scale contained 51 items that measured four dimensions of sensation seeking. The dimension ‘Thrill and adventure seeking’ contained 12 items, ‘Experience seeking’ had 14 items, ‘Boredom susceptibility’ contained 13 items and ‘Disinhibition’ was assessed with 12 items . Items were answered on a 5-point Likert-type scale ranging from totally disagree (1) to totally agree (5). Summed scores were computed for all four dimensions (after reversing negatively keyed items). If more than 2 items were missing per dimension, summed scores were not computed. If less than 2 items were missing per dimension, missingness was imputed with the average value on that dimension for a particular individual. In addition, an overall sensation seeking score was computed by adding the summed scores of the

four dimensions divided by the number of items (sss = tas/12 + es/14 + bs/13 + dis/12). Data on sensation seeking came from ANTR surveys 1991, 1993, 1997, 2000 and 2002.

Behavioral and emotional problems in children (YNTR, 3-12 years)

Parental ratings of behavioral and emotional problems in 3 to 12-year old children were collected by age appropriate versions of the Child Behavior Checklist, which are part of the Achenbach System of Empiral Based Assesment (www.aseba.org). Details of data collection are described elsewhere [33-35]. Parents report on the occurrence of behavioral problems in their children on a 3-point scale: 0 if the problem item was not true, 1 if the item was somewhat or sometimes true, and 2 if it was very true or often true. A syndrome scale is a summation of the corresponding items. At age 3, seven syndrome scales are obtained *(Emotionally Reactive, Anxious/Depressed, Somatic Complaints, Withdrawn, Overactive Behavior, Aggressive Behavior, Sleep)*, 2 broad groupings of syndromes (*Internalizing and Externalizing Problems)*, and the Total Problem score. At ages 7, 10, and 12 eight syndrome scales are obtained (*Anxious/ Depressed, Withdrawn, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior)* and 2 broad problems scales are obtained *(Internalizing and Externalizing Problems).*

3. Brain and Cognition

EEG was obtained in twins aged 5 through 71 years during 3–4 min of eyes-closed rest. After visual and semi-automated data cleaning, including removal of independent components linked to blink and heartbeat artifacts, EEG power was determined for the left frontal F3 electrode using Welch's periodogram method implemented in MATLAB. A full description of the sample and EEG registration/preprocessing can be found elsewhere (Smit et al., 2011).

Cognition data were obtained with age-appropriate psychometric IQ tests: RAKIT at ages 5, 7, and 10, the RAKIT, WISC at age 12, and WAIS at age 18. Data were collected in lab studies by trained research-assistants. Educational level (CITO scores) was assessed by standardized tests in 12-year olds during a 3 day testing period and these data were obtained from the parents and teachers of twins [36].

-*Educational attainment* (EA). In all ANTR surveys, participants were asked to indicate their highest educational level and whether or not they had completed this education with a diploma [37]. To obtain one measure of educational attainment for each participant, all adult data (age 21 years or older) for an individual were combined into a measure with 4 categories: 1. primary school; 2. lower vocational education and general secondary education; 3. intermediate vocational education and higher secondary education; 4. higher vocational education (HBO) and college/university education.

4. Growth and BMI

Birth weight: In YNTR, parents of twins reported birth weight in survey-1, which is collected after parents register their twins (age < 1 years) [38]. In the ANTR cohort, participants provided self-report birth weight in most surveys. In addition, parents reported on the birth weights of their offspring. All data available for an individual were combined and checked for consistency across time and rater. Birth weight was determined as the average of all valid data points, when the discrepancy was no larger than 200 grams [39].

Height: All ANTR surveys asked participants to provide their body height in cm. Data provided for the age of 18 years and above were checked for consistency across time. When the difference across time points did not exceed 5 cm, the height data were averaged to come to one adult body height [40]. In YNTR-surveys for 2 and 3 year olds), a parent was asked to report height and weight as measured by the routine health care program in the Netherlands (Youth Health

Services). In follow-up surveys (ages 5, 7, 10, 12 years) a parent was asked to report current weight and height. From age 14 onwards, the height and weights were self-reported by the twin.

-*BMI*: Body mass index (BMI) was obtained from maternal report in children and from self-report in adolescents and adults. BMI was calculated as $weight_{(kg)}$ / $height^2_{(m)}$. We used the body height measure as described above and weight from the most recent study in which both twins participated, or if this never was the case, the most recent study [41-42].

5. Metabolic risk factors and migraine

The metabolic risk factors were measured in blood samples of twins who participated in the NTR biobank study [43].

- *Blood pressure* (BP) data were collected in multiple studies. BP was measured using brachial cuff measurements on the non-dominant arm by an automated Dynamap 845 recorder in the laboratory [44] or by the Spacelabs 90207 ambulatory monitoring device [45]. The laboratory measurements were obtained with subjects sitting quietly for periods varying from 3 to 8 minutes. A mean of 3 to 6 BP measurements was taken. In the ambulatory BP study the mean of all evening measurements when subjects were seated calmly was taken as the resting value (mean number of measurements 3.8,). The BP values in all four studies were corrected for the use of antihypertensive medication [46-47]. If subjects were taking antihypertensive medication a correction of +14 mmHg for SBP and +10 mmHg for DBP was made for the laboratory studies. For the ambulatory study, the correction was drug class specific, but average values were close to +14 mmHg for SBP and +10 mmHg for DBP.

- *Lipids* (total cholesterol, HDL, LDL and triglycerides) were assessed after overnight fasting [43]. Total cholesterol HDL-cholesterol and triglyceride levels were measured in heparin plasma using the Vitros 250 total cholesterol assay, the Vitros 250 direct HDL cholesterol assay and the Vitros 250 Triglycerides assay (Johnson & Johnson, Rochester, USA). LDL-cholesterol was calculated using the Friedewald formula [48]. Subjects using lipid lowering medication at the

time of blood sampling were excluded from the analyses. For HDL, one extreme outlier was excluded from the analysis (HDL=6.56). For triglycerides an LN-transformation was applied .

-*glucose metabolism* (fasting glucose, fasting insulin and HbA1C) parameters were assessed after overnight fasting [43]). Blood glucose was measured using the Vitros 250 Glucose assay (Johnson & Johnson, Rochester, USA) and insulin was measured using the Immulite 1000 Insulin Method (Diagnostic Product Corporation, Los Angeles, USA). Hba1c was determined in EDTA whole blood using the Nyocard HbA1c assay (Axis-Shield, Oslo, Norway). For the analyses of glucose, insuline and all lipids, subjects who had not fasted from 24h the night before the home visit were excluded, and subjects taking medication for diabetes and subjects with a fasting glucose level >= 7 were excluded. For insuline an LN-transformation was applied .

-*Fibrinogen* was measured in a 4.5 ml CTAD tube that was stored during transport in melting ice. Fibrinogen levels were determined on a STA Compact Analyzer Diagnostica Stago, France), using STA Fibrinogen (Diagnostica Stago, France). For fibrinogen a cut off value of 5 pg/ml was used for exlusion of subjects.

-*C-reactive protein (CRP)* was measured in heparin plasma, CRP level was determined using the Immulite 1000 CRP assay (Diagnostic Product Corporation, USA). Data were excluded when values of CRP rose above 25 pg/ml,

-*Tumor necrosis factor-alpha (TNF-α), Interleukin-6 (IL-6)* and *Interleukin-receptor-6 (IL6R)* were measured in EDTA plasma, using R&D systems. IL6 and TNF- α data were excluded when values exceeded 15pg/ml and IL6-R data were excluded when exceeding 100.000 pg/ml.

-Data on the liver enzymes A*spartate Aminotransferase (AST)*, A*lanine Aminotransferase (ALT)* and G*amma-Glutamyl-Transferase (GGT)* were determined in plasma that was collected in heparin plasma tubes using Vitros assays were used (Willemsen et al., 2010). Liver enzyme levels were excluded for six individuals with liver disease (ICD-10 codes K70-K77), 12 individuals with extreme outliers (AST >79 U/L; ALT >63 U/L; GGT >296 U/L;

values >9 SD above the mean) and four twins without information on zygosity, resulting in a final data set of 3,662 individuals. To obtain more normally distributed variables, liver enzyme levels were ln-transformed and multiplied by 10 to avoid problems during estimation that result when analyzing very small covariances.

-*Migraine* was assessed with ANTR survey items that cover the symptoms included in the official diagnostic criteria for migraine (ICHD-II) as defined by the International Headache Society. The item variables were analyzed with latent class analysis [49], resulting in four groups of individuals, empirically classified based on their patterns of headache symptomatology, which differed primarily in the severity of their headaches. The two most severely affected classes, ("moderate/severe migraine" and "mild migrainous headache") were classified as affected. [50].

References

1. Vink, J.M., G. Willemsen, and D.I. Boomsma, *Heritability of smoking initiation and nicotine dependence.* Behavior Genetics, 2005. **35**(4): p. 397-406.
2. Vink, J.M., A.B. Smit, E.J.C. De Geus, P. Sullivan, G. Willemsen, J.J. Hottenga, J. Smit, W.J. Hoogendijk, F.G. Zitman, L. Peltonen, J. Kaprio, N.L. Pedersen, P. Magnusson, T.D. Spector, K.O. Kyvik, K.I. Morley, A.C. Heath, N.G. Martin, R.G.J. Westendorp, P.E. Slagboom, H. Tiemeier, A. Hofman, A. Uitterlinden, Y. Aulchenko, N. Amin, C. van Duijn, B.W. Penninx, and D.I. Boomsma, *Genome-wide association study of smoking initiation and current smoking.* The American Journal of Human Genetics, 2009. **84**: p. 1-13.
3. Heatherton, T.F., L.T. Kozlowski, R.C. Frecker, and K.O. Fagerstrom, *The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire.* British journal of addiction, 1991. **86**: p. 1119-1127.
4. Vink, J.M., G. Willemsen, A.L. Beem, and D.I. Boomsma, *The Fagerstrom Test for Nicotine Dependence in a Dutch sample of daily smokers and ex-smokers.* Addictive Behaviors, 2005. **30**(3): p. 575-579.
5. Vink, J.M., L. Wolters, M.C. Neale, and D.I. Boomsma, *Heritability of cannabis initiation in Dutch adult twins.* Addictive Behaviors, 2010. **35**(2): p. 172-174.
6. Distel, M.A., J.M. Vink, M. Bartels, C.E.M. van Beijsterveldt, M.C. Neale, and D.I. Boomsma, *Age moderates non-genetic influences on the initiation of cannabis use: a twin-sibling study in Dutch adolescents.* Addiction, 2011. **106**: p. 1658–1666.
7. Ewing, J.A., *Detecting alcoholism—the cage questionnaire.* J Am Med Assoc, 1984. **252**: p. 1905-1907.

8. Dhalla, S. and J.A. Kopec, *The CAGE questionnaire for alcohol misuse: a review of reliability and validity studies.* Clin Investig Med, 2007. **30**(33-41).

9. van Beek, J.H.D.A., K.S. Kendler, M.H.M. De Moor, L.M. Geels, M. Bartels, V. J.M., S.M. van den Berg, G. Willemsen, and D.I. Boomsma, *Stable Genetic Effects on Symptoms of Alcohol Abuse and Dependence from Adolescence into Early Adulthood.* Behavior Genetics, in press.

10. Geels, L.M., M. Bartels, C.E.M. van Beijsterveldt, G. Willemsen, N. van der Aa, J.M. Vink, and D.I. Boomsma, *Trends in adolescent alcohol use: Effects of age, sex and cohort.* Addiction, in press.

11. Vink, J.M., A.S. Staphorsius, and D.I. Boomsma, *A genetic analysis of coffee consumption in a sample of Dutch twins.* Twin Res Hum Genet., 2009. **12**(2): p. 127-31.

12. Ainsworth, B.E., W.L. Haskell, M.C. Whitt, M.L. Irwin, A.M. Swartz, S.J. Strath, W.L. O'Brien, D.R.J. Bassett, K.H. Schmitz, P.O. Emplaincourt, D.R.J. Jacobs, and A.S. Leon, *Compendium of physical activities: an update of activity codes and MET intensities.* Med Sci Sports Exerc., 2000. **32**(9 Suppl): p. S498-504.

13. Stubbe, J.H., D.I. Boomsma, J.M. Vink, B.K. Cornes, N.G. Martin, A. Skytthe, K.O. Kyvik, R.J. Rose, U.M. Kujala, J. Kaprio, J.R. Harris, N.L. Pedersen, J. Hunkin, T.D. Spector, and E.J.C. de Geus, *Genetic Influences on exercise participation in 37.051 twin pairs from seven countries.* PLoS ONE, 2006. **1**(2): p. e22.

14. Costa, P.T. and R.R. McCrae, *Professional manual: Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-factor-Inventory (NEO-FFI)*. 1992, Odessa, FL: Psychological Assesment Resources.

15. Hoekstra, H.A., J. Ormel, and F. De Fruyt, *Handleiding NEO persoonlijkheids-vragenlijsten NEO-PI-R en NEO-FFI.* , ed. S.T. Services. 1996, Lisse.

16. Marks, I.M. and A.M. Mathews, *Brief Standard Self-Rating for Phobic Patients.* Behav Res Ther 1979. **17**(263-267).

17. Sanavio, E., *Obsessions and compulsions: the Padua Inventory.* Behavior Research and Therapy, 1988. **26**(2): p. 169-177.

18. van Oppen, P., R.J. Hoekstra, and P.M. Emmelkamp, *The structure of obsessive-compulsive symptoms.* . Behavior Research and Therapy, 1995. **33**: p. 15-23.

19. Cath, D., D. van Grootheest, G. Willemsen, P. van Oppen, and D. Boomsma, *Environmental Factors in Obsessive-Compulsive Behavior: Evidence from Discordant and Concordant Monozygotic Twins.* Behavior Genetics, 2008. **38**(2): p. 108-120.

20. Morey, L.C., *The Personality Assessment Inventory: Professional manual.* 1991, Odessa, FL: Psychological Assessment Resources.

21. Distel, M.A., M.H.M. de Moor, and D.I. Boomsma, *Nederlandse vertaling van de Personality Assessment Inventory- borderline schaal (PAI-BOR): Normgegevens, factorstructuur en betrouwbaarheid.* Psychologie en Gezondheid, 2009. **37**(38-46).

22. Achenbach, T.M., *The Young Adult Self Report and Young Adult Behavior Checklist.* 1990, Burlington, VT: VT: University of Vermont, Department of Psychiatry.

23. Verhulst, F.C., J.v.d. Ende, and H.M. Koot, *Handleiding voor de Youth Self-Report*. 1997, Afdeling Kinder- en Jeugdpsychiatrie, Sophia Kinderziekenhuis / Academisch Ziekenhuis Rotterdam / Erasmus Universiteit Rotterdam: Rotterdam.

24. Spielberger, C.D., R.I. Gorsuch, and R.E. Lushene, *STAI Manual for the State-Trait Anxiety inventory.*, ed. P. Alto. 1970: CA: Consulting Psychologists Press.

25. Van der Ploeg, H., P.B. Defares, and C.D. Spielberger, *Zelfbeoordelingsvragenlijst STAI, versie DY-1 en DY-2*. 1979, Lisse: Swets & Zeitlinger.

26. Spielberger, C.D., G. Jacobs, J.S. Russell, and R.S. Crane, *Assessment of anger: the state-trait anger scale.* , in *Advances in Personality Assessment. Vol. 2.*, J.N.B.C.D. Spielberger, Editor. 1983, Erlbaum: Hillside, NJ.

27. van der Ploeg, H.M., P.B. Defares, and C.D. Spielberger, *Handleiding bij de zelf-analyse vragenlijst ZAV. Een vragenlijst voor het meten van boosheid en woede, als toestand en als episode. Een Nederlandse bewerking van de Spielberger State-Trait Anger Scale.* 1982, Lisse, the Netherlands: Swets & Zeitlinger.

28. Hughes, M.E., L.J. Waite, L.C. Hawkley, and J.T. Cacioppo, *A short scale for measuring loneliness in large surveys—results from two population-based studies.* Res Aging 2004. **26**: p. 655-672.

29. Russell, D., L.A. Peplau, and C.E. Cutrona, *The Revised Ucla Loneliness Scale—concurrent and discriminant validity evidence.* J Pers Soc Psychol, 1980. **39**: p. 472-480.

30. Zuckerman, M., *Dimensions of sensation seeking.* J Consult Clin Psychol, 1971. **36**: p. 45-52.

31. Feij, J.A., P.H. Dekker, J.R. Koopmans, and D.I. Boomsma, *Nieuwe normen en stabiliteitsgegevens voor de Spanningsbehoeftelijst (SBL).* Nederlands Tijdschrift voor de Psychologie, 1997. **52**: p. 131-134.

32. Feij, J.A. and R.W. Van Zuilen, *Handleiding bij de Spanningsbehoeftelijst (SBL)*. 1984, Lisse: Swets & Zeitlinger.

33. Bartels, M., C.E.M. Van Beijsterveldt, E.M. Derks, T.M. Stroet, T.J.C. Polderman, J.J. Hudziak, and D.I. Boomsma, *Young Netherlands Twin Register (Y-NTR): a longitudinal multiple informant study of problem behavior.* Twin Res Hum Genet., 2007. **10**(1): p. 3-11.

34. Bartels, M., E.J.C.G. Van den Oord, J.J. Hudziak, M.J.H. Rietveld, C.E.M. van Beijsterveldt, and D.I. Boomsma, *Genetic and environmental mechanisms underlying stability and change in problem behaviors at age 3, 7, 10 and 12.* Developmental Psychology, 2004. **40**(5): p. 852-867.

35. Derks, E.M., J.J. Hudziak, C.E.M. Beijsterveldt, C.V. Dolan, and D.I. Boomsma, *A Study of Genetic and Environmental Influences on Maternal and Paternal CBCL Syndrome Scores in a Large Sample of 3-Year-Old Dutch Twins.* Behavior Genetics, 2004. **34**(6): p. 571-583.

36. Bartels, M., M.J. Rietveld, G.C. Van Baal, and D.I. Boomsma, *Heritability of educational achievement in 12-year-olds and the overlap with cognitive ability.* Twin Res Hum Genet., 2002. **5**(6): p. 544-553.

37. Stronks, K., H. van de Mheen, J. van den Bos, and J. Mackenbach, *The interrelationship between income, health and employment status.* Int J Epidemiol., 1997. **26**(592-600).

38. Gielen, M., C.E.M. van Beijsterveldt, C. Derom, R. Vlietinck, J.G. Nijhuis, M.P.A. Zeegers, and D.I. Boomsma, *Secular trends in gestational age and birthweight in twins.* Human Reproduction, 2010. **25**(9): p. 2346-2353.

39. Freathy, R.M., D.O. Mook-Kanamori, U. Sovio, I. Prokopenko, N.J. Timpson, D.J. Berry, N.M. Warrington, E. Widen, J.J. Hottenga, M. Kaakinen, L.A. Lange, J.P. Bradfield, M. Kerkhof, J.A. Marsh, R. Magi, C.-M. Chen, H.N. Lyon, M. Kirin, L.S. Adair, Y.S. Aulchenko, A.J. Bennett, J.B. Borja, N. Bouatia-Naji, P. Charoen, L.J.M. Coin, D.L. Cousminer, E.J.C. de Geus, P. Deloukas, P. Elliott, D.M. Evans, P. Froguel, B. Glaser, C.J. Groves, A.-L. Hartikainen, N. Hassanali, J.N. Hirschhorn, A. Hofman, J.M.P. Holly, E. Hypponen, S. Kanoni, B.A. Knight, J. Laitinen, C.M. Lindgren, W.L. McArdle, P.F. O'Reilly, C.E. Pennell, D.S. Postma, A. Pouta, A.

Ramasamy, N.W. Rayner, S.M. Ring, F. Rivadeneira, B.M. Shields, D.P. Strachan, I. Surakka, A. Taanila, C. Tiesler, A.G. Uitterlinden, C.M. van Duijn, A.H. Wijga, G. Willemsen, H. Zhang, J. Zhao, J.F. Wilson, E.A.P. Steegers, A.T. Hattersley, J.G. Eriksson, L. Peltonen, K.L. Mohlke, S.F.A. Grant, H. Hakonarson, G.H. Koppelman, G.V. Dedoussis, J. Heinrich, M.W. Gillman, L.J. Palmer, T.M. Frayling, D.I. Boomsma, G.D. Smith, C. Power, V.W.V. Jaddoe, and M.-R. Jarvelin, *Variants in ADCY5 and near CCNL1 are associated with fetal growth and birth weight.* Nat Genet, 2010. **42**(5): p. 430-435.

40. Lango Allen, H., K. Estrada, G. Lettre, S.I. Berndt, M.N. Weedon, F. Rivadeneira, C.J. Willer, A.U. Jackson, S. Vedantam, S. Raychaudhuri, T. Ferreira, A.R. Wood, R.J. Weyant, A.V. Segre, E.K. Speliotes, E. Wheeler, N. Soranzo, J.-H. Park, J. Yang, D. Gudbjartsson, N.L. Heard-Costa, J.C. Randall, L. Qi, A. Vernon Smith, R. Magi, T. Pastinen, L. Liang, I.M. Heid, J.a. Luan, G. Thorleifsson, T.W. Winkler, M.E. Goddard, K. Sin Lo, C. Palmer, T. Workalemahu, Y.S. Aulchenko, A. Johansson, M. Carola Zillikens, M.F. Feitosa, T. Esko, T. Johnson, S. Ketkar, P. Kraft, M. Mangino, I. Prokopenko, D. Absher, E. Albrecht, F. Ernst, N.L. Glazer, C. Hayward, J.-J. Hottenga, K.B. Jacobs, J.W. Knowles, Z. Kutalik, K.L. Monda, O. Polasek, M. Preuss, N.W. Rayner, N.R. Robertson, V. Steinthorsdottir, J.P. Tyrer, B.F. Voight, F. Wiklund, J. Xu, J. Hua Zhao, D.R. Nyholt, N. Pellikka, M. Perola, J.R.B. Perry, I. Surakka, M.-L. Tammesoo, E.L. Altmaier, N. Amin, T. Aspelund, T. Bhangale, G. Boucher, D.I. Chasman, C. Chen, L. Coin, M.N. Cooper, A.L. Dixon, Q. Gibson, E. Grundberg, K. Hao, M. Juhani Junttila, L.M. Kaplan, J. Kettunen, I.R. Konig, T. Kwan, R.W. Lawrence, D.F. Levinson, M. Lorentzon, B. McKnight, A.P. Morris, M. Muller, J. Suh Ngwa, S. Purcell, S. Rafelt, R.M. Salem, E. Salvi, S. Sanna, J. Shi, U. Sovio, J.R. Thompson, M.C. Turchin, L. Vandenput, D.J. Verlaan, V. Vitart, C.C.

White, A. Ziegler, P. Almgren, A.J. Balmforth, H. Campbell, L. Citterio, A. De Grandi, A. Dominiczak, J. Duan, P. Elliott, R. Elosua, J.G. Eriksson, N.B. Freimer, E.J.C. Geus, N. Glorioso, S. Haiqing, A.-L. Hartikainen, A.S. Havulinna, A.A. Hicks, J. Hui, W. Igl, T. Illig, A. Jula, E. Kajantie, T.O. Kilpelainen, M. Koiranen, I. Kolcic, S. Koskinen, P. Kovacs, J. Laitinen, J. Liu, M.-L. Lokki, A. Marusic, A. Maschio, T. Meitinger, A. Mulas, G. Pare, A.N. Parker, J.F. Peden, A. Petersmann, I. Pichler, K.H. Pietilainen, A. Pouta, M. Ridderstrale, J.I. Rotter, J.G. Sambrook, A.R. Sanders, C. Oliver Schmidt, J. Sinisalo, J.H. Smit, H.M. Stringham, G. Bragi Walters, E. Widen, S.H. Wild, G. Willemsen, L. Zagato, L. Zgaga, P. Zitting, H. Alavere, M. Farrall, W.L. McArdle, M. Nelis, M.J. Peters, S. Ripatti, J.B.J. van Meurs, K.K. Aben, K.G. Ardlie, J.S. Beckmann, J.P. Beilby, R.N. Bergman, S. Bergmann, F.S. Collins, D. Cusi, M. den Heijer, G. Eiriksdottir, P.V. Gejman, A.S. Hall, A. Hamsten, H.V. Huikuri, C. Iribarren, M. Kahonen, J. Kaprio, S. Kathiresan, L. Kiemeney, T. Kocher, L.J. Launer, T. Lehtimaki, O. Melander, T.H. Mosley Jr, A.W. Musk, M.S. Nieminen, C.J. O/'Donnell, C. Ohlsson, B. Oostra, L.J. Palmer, O. Raitakari, P.M. Ridker, J.D. Rioux, A. Rissanen, C. Rivolta, H. Schunkert, A.R. Shuldiner, D.S. Siscovick, M. Stumvoll, A. Tonjes, J. Tuomilehto, G.-J. van Ommen, J. Viikari, A.C. Heath, N.G. Martin, G.W. Montgomery, M.A. Province, M. Kayser, A.M. Arnold, L.D. Atwood, E. Boerwinkle, S.J. Chanock, P. Deloukas, C. Gieger, H. Gronberg, P. Hall, A.T. Hattersley, C. Hengstenberg, W. Hoffman, G. Mark Lathrop, V. Salomaa, S. Schreiber, M. Uda, D. Waterworth, A.F. Wright, T.L. Assimes, I. Barroso, A. Hofman, K.L. Mohlke, D.I. Boomsma, M.J. Caulfield, L. Adrienne Cupples, J. Erdmann, C.S. Fox, V. Gudnason, U. Gyllensten, T.B. Harris, R.B. Hayes, M.-R. Jarvelin, V. Mooser, P.B. Munroe, W.H. Ouwehand, B.W. Penninx, P.P. Pramstaller, T. Quertermous, I. Rudan, N.J. Samani, T.D. Spector, H. Volzke, H. Watkins, J.F.

Wilson, L.C. Groop, T. Haritunians, F.B. Hu, R.C. Kaplan, A. Metspalu, K.E. North, D. Schlessinger, N.J. Wareham, D.J. Hunter, J.R. O/'Connell, D.P. Strachan, H.E. Wichmann, I.B. Borecki, C.M. van Duijn, E.E. Schadt, U. Thorsteinsdottir, L. Peltonen, A.G. Uitterlinden, P.M. Visscher, N. Chatterjee, R.J.F. Loos, M. Boehnke, M.I. McCarthy, E. Ingelsson, C.M. Lindgren, G.R. Abecasis, K. Stefansson, T.M. Frayling and J.N. Hirschhorn, *Hundreds of variants clustered in genomic loci and biological pathways affect human height.* Nature, 2010. **467**(7317): p. 832-838.

41. Speliotes, E.K., C.J. Willer, S.I. Berndt, K.L. Monda, G. Thorleifsson, A.U. Jackson, H.L. Allen, C.M. Lindgren, J.a. Luan, R. Magi, J.C. Randall, S. Vedantam, T.W. Winkler, L. Qi, T. Workalemahu, I.M. Heid, V. Steinthorsdottir, H.M. Stringham, M.N. Weedon, E. Wheeler, A.R. Wood, T. Ferreira, R.J. Weyant, A.V. Segre, K. Estrada, L. Liang, J. Nemesh, J.-H. Park, S. Gustafsson, T.O. Kilpelainen, J. Yang, N. Bouatia-Naji, T. Esko, M.F. Feitosa, Z. Kutalik, M. Mangino, S. Raychaudhuri, A. Scherag, A.V. Smith, R. Welch, J.H. Zhao, K.K. Aben, D.M. Absher, N. Amin, A.L. Dixon, E. Fisher, N.L. Glazer, M.E. Goddard, N.L. Heard-Costa, V. Hoesel, J.-J. Hottenga, A. Johansson, T. Johnson, S. Ketkar, C. Lamina, S. Li, M.F. Moffatt, R.H. Myers, N. Narisu, J.R.B. Perry, M.J. Peters, M. Preuss, S. Ripatti, F. Rivadeneira, C. Sandholt, L.J. Scott, N.J. Timpson, J.P. Tyrer, S. van Wingerden, R.M. Watanabe, C.C. White, F. Wiklund, C. Barlassina, D.I. Chasman, M.N. Cooper, J.-O. Jansson, R.W. Lawrence, N. Pellikka, I. Prokopenko, J. Shi, E. Thiering, H. Alavere, M.T.S. Alibrandi, P. Almgren, A.M. Arnold, T. Aspelund, L.D. Atwood, B. Balkau, A.J. Balmforth, A.J. Bennett, Y. Ben-Shlomo, R.N. Bergman, S. Bergmann, H. Biebermann, A.I.F. Blakemore, T. Boes, L.L. Bonnycastle, S.R. Bornstein, M.J. Brown, T.A. Buchanan, F. Busonero, H. Campbell, F.P. Cappuccio, C. Cavalcanti-Proenca, Y.-D.I. Chen, C.-M. Chen, P.S. Chines, R. Clarke, L. Coin, J. Connell,

I.N.M. Day, M.d. Heijer, J. Duan, S. Ebrahim, P. Elliott, R. Elosua, G. Eiriksdottir, M.R. Erdos, J.G. Eriksson, M.F. Facheris, S.B. Felix, P. Fischer-Posovszky, A.R. Folsom, N. Friedrich, N.B. Freimer, M. Fu, S. Gaget, P.V. Gejman, E.J.C. Geus, C. Gieger, A.P. Gjesing, A. Goel, P. Goyette, H. Grallert, J. Graszler, D.M. Greenawalt, C.J. Groves, V. Gudnason, C. Guiducci, A.-L. Hartikainen, N. Hassanali, A.S. Hall, A.S. Havulinna, C. Hayward, A.C. Heath, C. Hengstenberg, A.A. Hicks, A. Hinney, A. Hofman, G. Homuth, J. Hui, W. Igl, C. Iribarren, B. Isomaa, K.B. Jacobs, I. Jarick, E. Jewell, U. John, T. Jorgensen, P. Jousilahti, A. Jula, M. Kaakinen, E. Kajantie, L.M. Kaplan, S. Kathiresan, J. Kettunen, L. Kinnunen, J.W. Knowles, I. Kolcic, I.R. Konig, S. Koskinen, P. Kovacs, J. Kuusisto, P. Kraft, K. Kvaloy, J. Laitinen, O. Lantieri, C. Lanzani, L.J. Launer, C. Lecoeur, T. Lehtimaki, G. Lettre, J. Liu, M.-L. Lokki, M. Lorentzon, R.N. Luben, B. Ludwig, P. Manunta, D. Marek, M. Marre, N.G. Martin, W.L. McArdle, A. McCarthy, B. McKnight, T. Meitinger, O. Melander, D. Meyre, K. Midthjell, G.W. Montgomery, M.A. Morken, A.P. Morris, R. Mulic, J.S. Ngwa, M. Nelis, M.J. Neville, D.R. Nyholt, C.J. O'Donnell, S. O'Rahilly, K.K. Ong, B. Oostra, G. Pare, A.N. Parker, M. Perola, I. Pichler, K.H. Pietilainen, C.G.P. Platou, O. Polasek, A. Pouta, S. Rafelt, O. Raitakari, N.W. Rayner, M. Ridderstrale, W. Rief, A. Ruokonen, N.R. Robertson, P. Rzehak, V. Salomaa, A.R. Sanders, M.S. Sandhu, S. Sanna, J. Saramies, M.J. Savolainen, S. Scherag, S. Schipf, S. Schreiber, H. Schunkert, K. Silander, J. Sinisalo, D.S. Siscovick, J.H. Smit, N. Soranzo, U. Sovio, J. Stephens, I. Surakka, A.J. Swift, M.-L. Tammesoo, J.-C. Tardif, M. Teder-Laving, T.M. Teslovich, J.R. Thompson, B. Thomson, A. Tonjes, T. Tuomi, J.B.J. van Meurs, G.-J. van Ommen, V. Vatin, J. Viikari, S. Visvikis-Siest, V. Vitart, C.I.G. Vogel, B.F. Voight, L.L. Waite, H. Wallaschofski, G.B. Walters, E. Widen, S. Wiegand, S.H. Wild, G. Willemsen, D.R. Witte, J.C. Witteman, J. Xu, Q. Zhang, L. Zgaga, A.

Ziegler, P. Zitting, J.P. Beilby, I.S. Farooqi, J. Hebebrand, H.V. Huikuri, A.L. James, M. Kahonen, D.F. Levinson, F. Macciardi, M.S. Nieminen, C. Ohlsson, L.J. Palmer, P.M. Ridker, M. Stumvoll, J.S. Beckmann, H. Boeing, E. Boerwinkle, D.I. Boomsma, M.J. Caulfield, S.J. Chanock, F.S. Collins, L.A. Cupples, G.D. Smith, J. Erdmann, P. Froguel, H. Gronberg, U. Gyllensten, P. Hall, T. Hansen, T.B. Harris, A.T. Hattersley, R.B. Hayes, J. Heinrich, F.B. Hu, K. Hveem, T. Illig, M.-R. Jarvelin, J. Kaprio, F. Karpe, K.-T. Khaw, L.A. Kiemeney, H. Krude, M. Laakso, D.A. Lawlor, A. Metspalu, P.B. Munroe, W.H. Ouwehand, O. Pedersen, B.W. Penninx, A. Peters, P.P. Pramstaller, T. Quertermous, T. Reinehr, A. Rissanen, I. Rudan, N.J. Samani, P.E.H. Schwarz, A.R. Shuldiner, T.D. Spector, J. Tuomilehto, M. Uda, A. Uitterlinden, T.T. Valle, M. Wabitsch, G. Waeber, N.J. Wareham, H. Watkins, J.F. Wilson, A.F. Wright, M.C. Zillikens, N. Chatterjee, S.A. McCarroll, S. Purcell, E.E. Schadt, P.M. Visscher, T.L. Assimes, I.B. Borecki, P. Deloukas, C.S. Fox, L.C. Groop, T. Haritunians, D.J. Hunter, R.C. Kaplan, K.L. Mohlke, J.R. O'Connell, L. Peltonen, D. Schlessinger, D.P. Strachan, C.M. van Duijn, H.E. Wichmann, T.M. Frayling, U. Thorsteinsdottir, G.R. Abecasis, I. Barroso, M. Boehnke, K. Stefansson, K.E. North, M. I McCarthy, J.N. Hirschhorn, E. Ingelsson and R.J.F. Loos, *Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index.* Nat Genet, 2010. **42**(11): p. 937-948.

42. Schousboe, K., G. Willemsen, K.O. Kyvik, J. Mortensen, D.I. Boomsma, B.K. Cornes, C.J. Davis, C. Fagnani, J. Hjelmborg, J. Kaprio, M. de Lange, M. Luciano, N.G. Martin, N. Pedersen, Pietil, K.H. inen, A. Rissanen, S. Saarni, T.I.A. rensen, G.C.M. van Baal, and J.R. Harris, *Sex Differences in Heritability of BMI: A Comparative Study of Results from Twin Studies in Eight Countries.* Twin Research, 2003. **6**(5): p. 409-421.

43. Willemsen, G., E.J.C. de Geus, M. Bartels, C.E.M. van Beijsterveldt, A.I. Brooks, G.F. Estourgie-van Burk, D.A. Fugman, C. Hoekstra, J.J. Hottenga, K. Kluft, P. Meijer, G.W. Montgomery, P. Rizzu, D. Sondervan, A.B. Smit, S. Spijker, H.E.D. Suchiman, J.A. Tischfield, T. Lehner, P.E. Slagboom, and D.I. Boomsma, *The Netherlands Twin Register Biobank:A Resource for Genetic Epidemiological Studies.* Twin Res Hum Genet., 2010. **13**(3): p. 231-245.

44. Hottenga, J.J., D.I. Boomsma, H.M. Kupper, D. Posthuma, H. Snieder, A.H.M. Willemsen, and E.J.C. De Geus, *Heritability and stability of resting blood pressure.* Twin Res Hum Genet., 2005. **8**: p. 499-508.

45. Kupper, N., G. Willemsen, H. Riese, D. Posthuma, D.I. Boomsma, and E.J.C. de Geus, *Heritability of daytime ambulatory blood pressure in an extended twin design.* Hypertension, 2005. **45**: p. 80-85.

46. Cui, J.S., J.L. Hopper, and S.B. Harrap, *Antihypertensive Treatments Obscure Familial Contributions to Blood Pressure Variation.* Hypertension, 2003. **41**(2): p. 207-210.

47. Palmer, L.J., *Loosening the Cuff: important new advances in modeling antihypertensive treatment effects in genetic studies of hypertension.* Hypertension, 2003. **41**(2): p. 197-198.

48. Friedewald, W.T., R.I. Levy, and D.S. Frerickson, *Estimation of the concentration of low-density lipoprotein cholestrol in plasma, without use of the preparative ultracentrifuge.* . Clinical Chemistry, 1972. **18**: p. 499-502.

49. McCutcheon, A.L., *Latent Class Analysis*. 1987, Beverly Hills: Sage Publications.

50. Ligthart, L., D.R. Nyholt, B.W.J.H. Penninx, and D.I. Boomsma, *The Shared Genetics of Migraine and Anxious Depression.* Headache: The Journal of Head and Face Pain, 2010. **50**(10): p. 1549-1560.

Supplemental Table S1. Overview of available data for Lifestyle, including source (cohort), mean age, number of subjects, number of complete twin pairs, number of incomplete twins and percentage female participants. The ANTR surveys are part of the longitudinal study to health and personality of the Netherlands Twin Register: For Cohort/survey: 1 = ANTR data collected in 1991, 2= data collected in 1993, 3= data collected in 1995, 4- data collected in 1997, 5= data collected in 2000, 6= data collected in 2002, 7= data collected in 2004, 8= data collected in 2009, bb= bio bank project. For YNTR data the birth cohorts are given. Age (Range) = mean age of the sample and age range. N ss = Number of subjects, N cp = number of complete twin pairs, N icp = number of incomplete twins; Prev = prevalence, ß age = regression coefficient of age on mean/prevalence.

Table S1A. Lifestyle - Adults

Phenotype	Cohort/survey	Age (range)	N Ss	N cp	N icp	% ♀	Prev ♂	Prev ♀	ß age ♂	ß age ♀
Ever smoked	1,2,3,4,5,6,7,8	29.3 (18-65)	10004	4290	1424	64%	44%	40%	-.26	-.28
Current smoking	1,2,3,4,5,6,7,8	29.3 (18-65)	10004	4290	1424	64%	28%	22%	.05	.005
Nicotine depend	5,6,7,8	1971 (1940-1993)*	3360	852	1656	66%	38%	33%	.05	-.17
Ever cannabis	2, 3, 5, 8	28.7 (18 – 65)	7620	2641	2338	65%	34%	23%	.16	0.17
Regular drinking	7,8	32.0 (18-65)	7189	2540	2109	68%	11%	6%	.09	-0.005
Alcohol problems 0	2,3,4,5,6,7,8	1975 (1930-1992)*	11697	4703	2291	64%	61%	78%	.04	0.06
1							21%	13%	.07	0.11
2							11%	6%	.03	0.10
>2							7%	3%		
Coffee -0-2 cups	5	30.1 (18-65)	4287	1524	1239	67%	37%	63%	-.39	-.41
-3-5 cups							48%	21%	-.31	-.32
->6 cups							15%	16%		
Exercise	1,2,3,4,5,6,7,8	28.6 (18-65)	8605	3421	1763	65%	52%	49%	.102	0.052

*birth cohort instead of age because based on longitudinal data

Table 1B. Lifestyle - Adolescents

Phenotype	Cohort/survey	Age (range)	N Ss	N cp	N icp	% ♀	Prev ♂	Prev ♀	ß age ♂	ß age ♀
Ever smoked	1984-1994	15.7 (14-19)	6075	2817	441	56%	14%	15%	-.29	-.27
Current smoking	1984-1994	15.7 (14-19)	6075	2817	441	56%	7%	8%	-.36	-.29
Ever cannabis	1984-1994	16.2 (13-20)	6208	2826	556	56%	14%	13%	-.66	-.66
Ever alcohol	1984-1994	15.7 (14-19)	5917	2701	515	56%	78%	75%	-.67	-.63
Early alcohol initiat	1984-1994	14.2 (13-15)	3630	1649	332	56%	75%	72%	-.08	-.14
Weekly alcohohol	1984-1994	15.7 (13-18)	5917	2701	515	56%	24%	17%	-.77	-.55

Supplemental S2. Overview of available data for Emotional and Behavioral Problems, including source (cohort), mean age, number of subjects, number of complete twin pairs, number of incomplete twins and percentage female participants. The ANTR surveys are part of the longitudinal study to health and personality of the Netherlands Twin Register: For Cohort/survey: 1 = ANTR data collected in 1991, 2= data collected in 1993, 3= data collected in 1995, 4- data collected in 1997, 5= data collected in 2000, 6= data collected in 2002, 7= data collected in 2004, 8= data collected in 2009, bb= bio bank project. For YNTR data the birth cohorts are given. Age (Range) = mean age of the sample and age range. N ss = Number of subjects, N cp = number of complete twin pairs, N icp = number of incomplete twins; Prev = prevalence, ß age = regression coefficient of age on mean/prevalence.

Table 2A. Emotional and Behavioral problems – Psychiatry adults

Phenotype	**Cohort/survey**	**Age (range)**	**N Ss**	**N cp**	**N icp**	**% ♀**	**Mean (var) ♂**		**Mean (var) ♀**		**ß age ♂**	**ß age ♀**
Anxious Dep ASR	1,3,4,5,6	29.5 (18-65)	7409	3238	933	63%	17.13	9.97	22.30	9.60	-.16	-.54
Anxiety	1,2,4,5,6	29.6 (18-65)	7463	3261	941	63%	32.08	9.17	35.23	7.99	-.21	-.52
OCS	6,8	31.7 (18-65)	7128	2499	2130	68%	7.38	5.41	7.20	5.36	-.56	-.66
Borderline	7,8	31.9 (18-65)	7085	2508	2069	69%	3.71	1.02	3.90	1.06	-.15	-.19
Anger	1,2,7	30.71 (18-65)	5744	2438	868	64%	15.60	4.49	15.74	4.23	-.58	-.68
ADHD	7,8	31.2 (18-65)	6327	2045	2237	68%	8.03	4.19	8.14	3.85	-.41	-.41
Phenotype	**Cohort/survey**	**Age (range)**	**N Ss**	**N cp**	**N icp**	**% ♀**	**Prev ♂**		**Prev ♀**		**ß age ♂**	**ß age ♀**
Social Fear	4, 5	29.0 (18-65)	5170	2053	1064	66%	62%		44%		.09	.06
Blood Fear	4,5	29.0 (18-65)	5170	2053	1064	66%	50%		53%		.09	.12
Loneliness	7,8	31.0 (18-65)	8380	3079	2222	68%	52%		45%		.067	.094
							22%		29%		.047	.045
							8%		12%		.007	.045
							18%		14%			

Cohort: 1 = 1991, 2=1993, 3=1995, 4-1997, 5=2000, 6=2002, 7=2004, 8=2009

Table 2B. Emotional and Behavioral problems - Personality adults

Phenotype	Cohort/survey	Age (range)	N Ss	N cp	N icp	% ♀	Mean (var) ♂		Mean (var) ♀		ß age ♂	ß age ♀
Extraversion	7	34.1 (18-65)	4731	1624	1483	70%	27.64	7.07	30.96	7.56	-.19	-.46
Neuroticism	7	34.1 (18-65)	4731	1624	1483	70%	42.67	5.61	42.29	5.88	-1.00	-.65
Conscientiousness	7	34.1 (18-65)	4731	1624	1483	70%	45.04	5.15	45.44	5.51	.28	-.02
Agreeableness	7	34.1 (18-65)	4731	1624	1483	70%	43.26	4.60	45.88	4.81	.74	.29
Openness to Experience	7	34.1 (18-65)	4731	1624	1483	70%	36.68	5.65	37.07	5.91	-.19	-.18
Thrill & adventure Seeking	1,2,4,5,6	28.5 (18-65)	7222	3108	1006	63%	39.20	9.63	32.59	9.63	-3.39	-4.37
Disinhibition	1,2,4,5,6	28.5 (18-65)	7222	3108	1006	63%	34.87	7.76	28.91	7.29	-1.86	-2.03
Experience seeking	1,2,4,5,6	28.5 (18-65)	7222	3108	1006	63%	35.10	8.22	32.64	8.09	-0.43	-0.87
Boredom susceptibility	1,2,4,5,6	28.5 (18-65)	7222	3108	1006	63%	37.17	7.21	35.01	7.47	-.79	-.83
Sensation seeking	1,2,4,5,6	28.5 (18-65)	7222	3108	1006	63%	11.54	1.88	10.15	1.82	-.53	-.66

Cohort: 1 = 1991, 2=1993, 3=1995, 4-1997, 5=2000, 6=2002, 7=2004, 8=2009

Table 2C. Emotional and Behavioral problems – Internalizing children

Phenotype	Cohort/survey	Age (range)	N Ss	N cp	N icp	% ♀	Mean (var) ♂		Mean (var) ♀		ß age ♂	ß age ♀
Anxious Depr – 3	1986-2004	3.3 (2.4-5.0)	31864	15873	118	50%	3.49	2.85	3.33	2.85	NA	NA
Anxious Depr -7	1986 - 99	7.4 (6.1-9.8)	10181	10106	75	50%	2.02	2.54	2.14	2.47	NA	NA
Anxious Depr -10	1986-1997	10.1 (8.7-12.9)	6970	6927	43	48%	2.30	2.94	2.41	2.79	NA	NA
Anxious Depr -12	1986-1998	12.2 (11.2-14.3)	6554	6516	38	49%	2.02	2.81	2.18	2.81	NA	NA
Internalizing -3	1986-2004	3.3 (2.4-5.0)	31864	15873	118	50%	4.81	3.66	4.43	3.84	NA	NA
Internalizing – 7	1986 - 99	7.4 (6.1-9.8)	10181	10106	75	49%	4.18	4.23	4.44	4.15	NA	NA
Internalizing -10	1986-1997	10.1 (8.7-12.9)	6970	6927	43	48%	4.45	4.74	4.69	4.50	NA	NA
Internalizing -12	1986-1998	12.2 (11.2-14.3)	6554	6516	38	49%	3.94	4.61	4.17	4.61	NA	NA
Somatic compl- 3	1986-2004	3.3 (2.4-5.0)	31864	15873	118	50%	0.63	3.74	0.70	3.66	NA	NA
Somatic compl - 7	1986 - 99	7.4 (6.1-9.8)	10181	10106	75	49%	0.76	1.29	0.89	1.17	NA	NA
Somatic compl - 10	1986-1997	10.1 (8.7-12.9)	6970	6927	43	48%	0.76	0.14	0.94	1.21	NA	NA
Somatic compl - 12	1986-1998	12.2 (11.2-14.3)	6554	6516	38	49%	0.66	1.26	0.81	1.17	NA	NA
Withdrawn - 3	1986-2004	3.3 (2.4-5.0)	31864	15873	118	50%	1.23	1.31	1.08	1.48	NA	NA
Withdrawn - 7	1986 - 99	7.4 (6.1-9.8)	10181	10106	75	49%	1.50	1.52	1.51	1.60	NA	NA
Withdrawn - 10	1986-1997	10.1 (8.7-12.9)	6970	6927	43	48%	1.51	1.65	1.46	1.71	NA	NA
Withdrawn - 12	1986-1998	12.2 (11.2-14.3)	6554	6516	38	49%	1.35	1.62	1.27	1.75	NA	NA

Table 2D Emotional and Behavioral problems – Externalizing children

Phenotype	Cohort/survey	Age (range)	N Ss	N cp	N icp	% ♀	Mean (var) ♂		Mean (var) ♀		ß age ♂	ß age ♀
Externalizing -3	1986-2004	3.3 (2.4-5.0)	31864	15873	118	50%	16.4	8.88	14.6	9.69	NA	NA
Externalizing -7	1986 - 99	7.4 (6.1-9.8)	10181	10106	75	49%	8.36	6.12	5.59	6.75	NA	NA
Externalizing -10	1986-1997	10.1 (8.7-12.9)	6970	6927	43	48%	7.60	5.38	5.40	6.86	NA	NA
Externalizing -12	1986-1998	12.2 (11.2-14.3)	6554	6516	38	49%	6.22	4.50	4.88	6.26	NA	NA
Aggression -3	1986-2004	3.3 (2.4-5.0)	31864	15873	118	50%	3.70	2.50	2.17	2.75	NA	NA
Aggression -7	1986 - 99	7.4 (6.1-9.8)	10181	10106	75	49%	7.12	5.22	4.73	5.63	NA	NA
Aggression -10	1986-1997	10.1 (8.7-12.9)	6970	6927	43	48%	6.42	4.59	4.51	5.62	NA	NA
Aggression - 12	1986-1998	12.2 (11.2-14.3)	6554	6516	38	49%	5.24	3.85	4.08	5.11	NA	NA
Opposition def - 3	1986-2004	3.3 (2.4-5.0)	31864	15873	118	50%	9.90	9.75	5.99	6.14	NA	NA
Rule-breaking beh - 7	1986 - 99	7.4 (6.1-9.8)	10181	10106	75	49%	1.34	1.00	1.87	1.53	NA	NA
Rule-breaking beh – 10	1986-1997	10.1 (8.7-12.9)	6970	6927	43	48%	1.21	0.82	1.23	1.62	NA	NA
Rule-breaking beh- 12	1986-1998	12.2 (11.2-14.3)	6554	6516	38	49%	1.03	0.67	1.15	1.59	NA	NA

Table 2E. Emotional and Behavioral problems – other scales children

Phenotype	Cohort/survey	Age (range)	N Ss	N cp	N icp	% ♀	Mean (var) ♂		Mean (var) ♀		ß age ♂	ß age ♀
Social problems - 7	1986 - 99	7.4 (6.1-9.8)	10181	10106	75	49%	1.32	1.48	1.10	1.61	NA	NA
Social problems - 10	1986-1997	10.1 (8.7-12.9)	6970	6927	43	48%	1.43	1.66	1.00	1.53	NA	NA
Social problems - 12	1986-1998	12.2 (11.2-14.3)	6554	6516	38	49%	1.35	1.57	0.82	1.62	NA	NA
Thought problems - 7	1986 - 99	7.4 (6.1-9.8)	10181	10106	75	49%	0.47	1.05	0.67	1.59	NA	NA
Thought problems - 10	1986-1997	10.1 (8.7-12.9)	6970	6927	43	48%	0.43	0.82	1.10	1.61	NA	NA
Thought problems - 12	1986-1998	12.2 (11.2-14.3)	6554	6516	38	49%	0.38	0.73	1.22	1.81	NA	NA
Sleep problems - 3	1986-2004	3.3 (2.4	31864	15873	118	50%	2.04	2.64	1.06	1.82	NA	NA
Total problems - 3	1986-2004	3.3 (2.4	31864	15873	118	50%	32.41	16.38	19.08	17.74	NA	NA
Overactive -3	1986-2004	3.3 (2.4-5.0)	31864	15873	118	50%	2.88	2.00	0.33	0.95	NA	NA
Attention problems -7	1986 - 99	7.4 (6.1-9.8)	10181	10106	75	49%	3.24	2.42	0.27	0.95	NA	NA
Attention problems- 10	1986-1997	10.1 (8.7-12.9)	6970	6927	43	48%	3.29	2.45	2.03	2.64	NA	NA
Attention problems-12	1986-1998	12.2 (11.2-14.3)	6554	6516	38	49%	2.97	2.25	1.98	2.87	NA	NA

Table S3. Brain and Cognition. Overview of available data, including source (cohort), mean age, number of subjects, number of complete twin pairs, number of incomplete twins and percentage female participants. The ANTR surveys are part of the longitudinal study to health and personality of the Netherlands Twin Register: For Cohort/survey: 1 = ANTR data collected in 1991, 2= data collected in 1993, 3= data collected in 1995, 4- data collected in 1997, 5= data collected in 2000, 6= data collected in 2002, 7= data collected in 2004, 8= data collected in 2009, bb= bio bank project. For YNTR data the birth cohorts are given. Age (Range) = mean age of the sample and age range. N ss = Number of subjects, N cp = number of complete twin pairs, N icp = number of incomplete twins; Prev = prevalence, ß age = regression coefficient of age on mean/prevalence.

Phenotype	Cohort/survey	Age (range)	N Sub-jects	N comp pairs	N incomp	% ♀	Mean (var) ♂		Mean (var) ♀		ß age ♂	ß age ♀
F3 power	1986-1988	5.3 (4.9 – 5.9)	366	165	51	51%	11.0	2.51	11.7	2.27	.25	-.04
F3 power	1986-1988	6.8 (6.4 – 7.5)	378	186	49	49%	11.5	2.46	12.1	2.49	.12	-.13
F3 power	1974-1977	16.1 (14.8-18.0)	426	213	54	54%	8.2	3.58	8.5	3.36	-.35	-.61
F3 power	1974-1977	17.6 (16.4-19.5)	387	191	53	53%	7.5	3.49	8.8	3.39	-.32	.60
F3 power	1955-1986	25.7 (18.7-33.9)	393	131	54	54%	5.3	3.60	6.8	4.15	-.33	-.56
F3 power	1929-1955	49.3 (36.0 – 71.0)	349	119	61	61%	3.7	3.68	5.6	4.05	-1.29	-.05
Educational level 12	1986-1996	11.92 (10.5-13.5)	8790	4074	642	53%	538.0	8.62	537.0	8.78	NA	NA
Educational attainment	1,2,3,4,5,6,7,8	32.23 (21.0-65.0)	7342	2874	1594	65%	4.90	1.57	4.90	1.50	.27	-.09
Educational attainment	1,2,3,4,5,6,7,8	31.92 (21.0-65.0)	6952	2571	1810	65%	109.4	14.5	109.5	14.1	NA	NA
IQ age 5	1986-88; 90-91	5.5 (4.9 -5.9)	889	444	1	50%	103.2	14.6	102.6	14.7	NA	NA
IQ age 7	1986-88	6.8 (6.4-7.5)	382	191	0	50%	108.3	14.6	105.6	16.1	NA	NA
IQ age 10	1986-88	10.0 (9.1-10.9)	392	196	0	50%	98.3	14.1	101.3	13.7	NA	NA
IQ age 12	1986-88; 90-91	12.2 (11.9-13.1)	734	366	2	52%	110.0	14.6	108.6	14.8	NA	NA
IQ age 18	1986-88; 76-77	17.9 (16.4-19.5)	758	375	8	53%	11.0	2.51	11.7	2.27	.25	-.04

Cohort: 1 = 1991, 2=1993, 3=1995, 4-1997, 5=2000, 6=2002, 7=2004, 8=2009, Age xx represents age-cohorts from the Y-NTR

Supplemental Table S4. Overview of available data for BMI and Height, including source (cohort), mean age, number of subjects, number of complete twin pairs, number of incomplete twins and percentage female participants. The ANTR surveys are part of the longitudinal study to health and personality of the Netherlands Twin Register: For Cohort/survey: 1 = ANTR data collected in 1991, 2= data collected in 1993, 3= data collected in 1995, 4- data collected in 1997, 5= data collected in 2000, 6= data collected in 2002, 7= data collected in 2004, 8= data collected in 2009, bb= bio bank project. For YNTR data the birth cohorts are given. Age (Range) = mean age of the sample and age range. N ss = Number of subjects, N cp = number of complete twin pairs, N icp = number of incomplete twins; Prev = prevalence, ß age = regression coefficient of age on mean/prevalence.

Table 4A. BMI and Height - adults

Phenotype	Cohort/survey	Age (range)	N Ss	N cp	N icp	% ♀	Mean (var) ♂		Mean (var) ♀		ß age ♂	ß age ♀
BMI adults	1,2,3,4,5,6,7,8	28.77 (18-65)	11370	4725	1920	64%	22.95	2.86	22.67	3.48	1.36	1.07
Height	1,2,3,4,5,6,7,8, bb	28.77 (18-65)	11340	4710	1920	64%	182.8	7.16	169.5	6.43	.87	.83

Cohort: 1 = 1991, 2=1993, 3=1995, 4-1997, 5=2000, 6=2002, 7=2004, 8=2009, bb=biobank

Table 4B. Birthweight, BMI and height - children and adolescents

Phenotype	Cohort/survey	Age (range)	N Ss	N cp	N icp	% ♀	Mean (var) ♂		Mean (var) ♀		ß age ♂	ß age ♀
Birth weight	1986-2004	36.6 (24-42)*	34617	17277	63	50%	2557	385	2451	369	399	385
BMI age 3	1986-2002	3.0 (2.5-3.5)	20828	10360	108	50%	15.74	1.23	15.48	1.31	-.11	-.06
BMI age 5	1986-2000	5.1 (4-6)	9841	4896	78	50%	15.05	1.24	14.95	1.53	-.04	-.07
BMI age 7	1986-1998	7.4 (6-8)	12154	6067	49	51%	15.28	1.65	15.38	1.92	.05	.15
BMI age 10	1986-1995	10.1 (9-11)	8764	4366	20	51%	16.30	2.07	16.52	2.28	.00	.13
BMI age 12	1985-1993	12.1 (11-13)	6234	3099	36	51%	17.21	2.35	17.51	2.53	.23	.17
BMI age 14	1990-1994	14.7 (13.9-16)	3275	1549	177	55%	18.96	2.30	19.41	2.62	.37	.35

BMI age 16	1988-1992	16.8 (15-18)	2256	1085	105	57%	20.30	2.37	20.66	2.77	.20	.13
Height age 1	1986-2003	1 (.6-1.3)	31829	15761	307	50%	61.63	2.77	59.75	2.74	13.69	14.08
Height age 2	1986-2003	2 (1.6-2.3)	23147	11470	207	50%	87.62	3.40	86.50	3.34	1.58	1.64
Height age 3	1986-2002	3 (2.5-3.5)	21120	10521	78	50%	97.28	3.82	96.34	3.83	1.50	1.52
Height age 5	1986-2000	5.1 (4-6)	10187	5079	29	50%	111.04	4.87	110.33	4.87	6.36	6.40
Height age 7	1986-1998	7.4 (6-8)	12331	6160	11	51%	128.56	5.70	127.85	5.66	2.25	2.42
Height age 10	1986-1995	10.1 (9-11)	8917	4448	19	51%	143.97	6.54	143.60	6.75	1.89	2.16
Height age 12	1985-1993	12.1 (11-13)	6344	3161	21	51%	154.91	7.38	155.95	7.35	1.94	2.00
Height age 14	1990-1994	14.7 (13.9-16)	3374	1630	114	55%	172.79	8.32	167.13	6.41	2.5	.74
Height age 16	1988-1992	16.8 (15-18)	2358	1137	83	57%	181.24	7.08	169.59	6.40	.47	.27

Table S5. Cardiovascular, metabolic and migraine. Overview of available data, including source (cohort), mean age, number of subjects, number of complete twin pairs, number of incomplete twins and percentage female participants. The ANTR surveys are part of the longitudinal study to health and personality of the Netherlands Twin Register: For Cohort/survey: 1 = ANTR data collected in 1991, 2= data collected in 1993, 3= data collected in 1995, 4- data collected in 1997, 5= data collected in 2000, 6= data collected in 2002, 7= data collected in 2004, 8= data collected in 2009, bb= bio bank project. For YNTR data the birth cohorts are given. Age (Range) = mean age of the sample and age range. N ss = Number of subjects, N cp = number of complete twin pairs, N icp = number of incomplete twins; Prev = prevalence, ß age = regression coefficient of age on mean/prevalence.

Phenotype	**Cohort/survey**	**Age (range)**	**N Ss**	**N cp**	**N icp**	% ♀	**Mean (var) ♂**		**Mean (var) ♀**		**ß age ♂**	**ß age ♀**
Systolic bp	1918-1998	26.6 (13-75)	2102	982	138	58%	127.1	11.2	120.4	11.4	.38	.77
Diastolic bp	1918-1998	26.6 (13-75)	2102	982	138	58%	72.5	8.6	71.7	8.46	3.99	1.52
HDL	bb	34.1 (18- 65)	3597	1278	1041	66%	1.24	0.30	1.51	0.37	-.02	.03
LDL	bb	34.1 (18 – 65)	3596	1279	1038	66%	2.96	0.84	2.82	0.84	.35	.32
Total cholesterol	bb	34.1 (18 – 65)	3600	1280	1040	66%	4.84	0.94	4.84	0.92	.43	.38
Triglycerides (ln)*	bb	34.1 (18– 65)	3600	1280	1040	66%	1.40	0.90	1.12	0.59	.14	.05
Glucose	bb	34.3 (18-65)	3595	1279	1037	66%	5.36	.50	5.16	.47	.09	.13
CRP (ln)*	bb	34.4 (18- 65)	3791	1398	995	65%	1.81	2.30	2.77	3.04	.23	-.02
TNFalpha (ln)*	bb	34.4 (18- 65)	3702	1400	902	66%	1.01	.85	1.09	1.18	.04	.03
IL6 (ln)*	bb	34.4(18- 65)	3708	1403	902	66%	1.40	1.46	1.40	1.33	.16	.11
IL6 receptor (ln)*	bb	34.4 (18-65)	3723	1415	893	66%	41963	11325	40438	11580	.01	.04
Insulin (ln)*	bb	34.2(18- 65)	3507	1224	1059	66%	8.84	5.79	8.43	5.49	.09	-.03
HbA1C	bb	34.2 (18-65)	3780	1405	970	65%	5.20	.44	5.24	.51	.03	.06
Liver enzyme: ALT	bb	35.6 (18-65)	3519	1266	987	65%	24.46	4.40	20.92	4.30	-.21	.36
Liver enzyme: GGT	bb	35.6 (18-65	3660	1365	930	66%	34.12	4.77	30.44	3.92	1.29	0.66
Phenotype	**Cohort/survey**	**Age (range)**	**N Ss**	**N cp**	**N icp**	% ♀	**Prev ♂**		**Prev ♀**		**ß**	**ß age**

									age ♂	♀
Migraine	6,7	33.5 (18-65)	5418	2010	1398	68%	14%	36%	-0.08	-0.03

* variable was ln transformed in analyses, the mean and variances of the original (untransformed) data are shown in this table.

Supplemental Table S6. Number of twin pairs required to detect significant difference between correlation in DZ same sex (DZss) twin pairs and DZ opposite sex (DZos) twin pairs.

		Power of the test at the significance level of 0.05 (with 1df)				
RDzss/Rdzos		**.75**	**.80**	**.90**	**.95**	**.99**
	.55/.225	212	240	321	397	561
	.50/.25	284	321	430	532	752
	.45/.225	383	433	580	718	1012
	.40/.20	525	594	795	984	1390
	.35/.175	735	831	1112	1375	1944
	.30/.15	1060	1198	1604	1984	2805
	.25/.125	1602	1812	2425	2999	4241
	.20/.10	2603	2944	3941	4875	6892
	.15/.075	4771	5396	7223	8933	12630

Supplementary Table S7A. Twin correlations for traits with significantly lower DZ opposite-sex correlations than DZ same-sex correlations. For these traits, a genetic model was fitted to the data to test whether the difference was due to different genes being expressed in men and women or environmental factors being less correlated in opposite-sex pairs (see Supplementary Table 11C and D for results).

	Model 1: 5 correlations (R)					**Best model**		
	Rmzm	Rdzm	Rmzf	Rdzf	Rdos	Rdzm	Rdzf	Rdos
Lifestyle -Adolescents								
Ever alcohol 14-18 year	0.82	0.69	0.91	0.77	0.61	0.73	0.73	**0.61**
Weekly alcohol use 14-18 year	0.70	0.64	0.74	0.66	0.46	0.65	0.65	**0.46**
Personality and Psychopathology – Personality adults								
Thrill & adventure adults	0.67	0.4	0.61	0.38	0.23	0.39	0.39	**0.23**
Sensation seeking adults	0.62	0.45	0.61	0.39	0.29	0.41	0.41	**0.29**
BMI and Height – children and adolescents								
Height -7	0.94	0.6	0.93	0.58	0.54	0.59	0.59	**0.54**
Metabolic risk factors & Migraine								
HDL	0.70	0.51	0.62	0.36	0.21	0.40	0.40	**0.21**

Supplementary Table S7B. Full model including additive genetic, common and unique environmental factors (a^2 , c^2 , and e^2 give explained variance for traits with evidence for qualitative sex differences; γ and ϕ represent respectively the genetic correlation and environmental correlation in DZ-opposite sex twin pairs

	a^2 M	a^2 F	c^2 M	c^2 F	e^2 M	e^2 F	$R_g$dz-os (γ)	$R_c$dz-os (φ)
Lifestyle -Adolescents								
Ever alcohol 14-18	0.27	0.27	0.56	0.63	0.18	0.09		0.80
Weekly alc 14-18	0.12	0.14	0.58	0.59	0.30	0.27		0.67
Personality and Psychopathology – Personality adults								
Thrill & adventure	0.54	0.47	0.13	0.15	0.33	0.39	0.18	
Sensation seeking	0.35	0.45	0.27	0.16	0.38	0.39	0.19	
Height								
Height -7 years	0.68	0.70	0.26	0.23	0.06	0.07	0.43	
Metabolic risk factors								
HDL	0.38	0.54	0.32	0.09	0.29	0.38	0.08	

Supplementary Table S7C. Parameter estimates based on most parsimonious model

	a^2 M	a^2 F	c^2 M	c^2 F	e^2 M	e^2 F	$R_g$dos (γ)	$R_c$dos (φ)
Lifestyle -Adolescents								
Ever alcohol 14-18	0.28		0.59		0.12			0.79
Weekly alc 14-18	0		0.69		0.31			0.66
Personality and Psychopathology – Personality adults								
Thrill & adventure	0.64		0		0.36		0.36	
Sensation seeking	0.41		0.20		0.39		0.20	
Height								
Height -7 years	0.69		0.25		0.06		0.43	
Metabolic risk factors								
HDL	0.70	0.62	0	0	0.30	0.38	0.50	

Supplementary Table S7D. Overview of the twin correlations for traits where DZ correlations were significantly different from each other (but Rdos not lower than Rdzm/Rdzf). No additional models were fitted.

	Model 1					**Best model**		
	Rmzm	Rdzm	Rmzf	Rdzf	Rdos	Rdzm	Rdzf	Rdos
Personality and Psychopathology – Internalizing behavior children								
Anxious Depr – 3	0.73	0.33	0.72	0.32	0.38	0.33	0.32	**0.38**
Anxious Depr -7	0.67	0.32	0.66	0.3	0.39	0.31	0.31	**0.39**
Anxious Depr -10	0.66	0.27	0.62	0.32	0.38	0.29	0.29	**0.38**
Anxious Depr -12	0.64	0.24	0.64	0.33	0.42	**0.24**	**0.32**	**0.42**
Internalizing -3	0.77	0.43	0.76	0.42	0.46	0.42	0.42	**0.46**
Internalizing – 7	0.73	0.41	0.73	0.45	0.51	0.43	0.43	**0.51**
Internalizing -10	0.66	0.38	0.7	0.44	0.49	0.41	0.41	**0.49**
Internalizing -12	0.73	0.38	0.7	0.45	0.52	0.42	0.42	**0.52**
Somatic complaints, CBCL - 3	0.93	0.88	0.95	0.93	0.93	**0.88**	**0.93**	**0.93**
Somatic complaints, CBCL - 7	0.54	0.28	0.63	0.41	0.37	**0.28**	**0.41**	**0.37**
Somatic complaints, CBCL - 10	0.53	0.23	0.63	0.34	0.33	**0.23**	**0.34**	**0.33**
Somatic complaints, CBCL - 12	0.52	0.36	0.63	0.42	0.44	0.39	0.39	**0.44**
Withdrawn, CBCL - 3	0.73	0.41	0.76	0.46	0.48	**0.41**	**0.46**	**0.48**
Withdrawn, CBCL - 7	0.69	0.26	0.67	0.33	0.35	**0.26**	**0.33**	**0.35**
Withdrawn, CBCL - 10	0.67	0.33	0.67	0.33	0.39	0.33	0.33	**0.38**
Withdrawn, CBCL - 12	0.7	0.32	0.65	0.3	0.42	0.31	0.31	**0.42**
Personality and Psychopathology – Externalizing behavior children								
Externalizing -12	0.85	0.46	0.81	0.48	0.55	0.47	0.47	**0.55**
Aggress beh - 12	0.83	0.44	0.79	0.45	0.51	0.44	0.44	**0.51**
Overactive -3	0.71	0.15	0.69	0.17	0.24	0.16	0.16	**0.24**
Rule-breaking behavior, CBCL - 12	0.79	0.49	0.83	0.53	0.63	0.51	0.51	**0.62**
Personality and Psychopathology – Other								
Social problems, CBCL - 7	0.76	0.36	0.74	0.29	0.34	**0.36**	**0.29**	**0.34**
Social problems, CBCL - 12	0.75	0.21	0.72	0.24	0.3	0.28	0.28	**0.33**
Thought problems, CBC - 10	0.58	0.18	0.45	0.2	0.26	0.19	0.19	**0.26**
Brain and cognition								
Education adults	0.69	0.38	0.68	0.53	0.39	**0.38**	**0.53**	**0.39**
BMI and height – children and adolescents								
BMI – 14	0.82	0.35	0.83	0.49	0.35	**0.35**	**0.49**	**0.35**

Supplemental Table S8. Overview of all published GWA studies for Height based on the database published on www.genome.gov/gwastudies (February 2011) and literature search in pub med. We selected this phenotypes because it is representative selection among all complex traits and it meets the criteria proposed by Visscher et al: at least three GWAS papers published in journals with a 2010-2011 journal impact factor >9 and at least one paper containing 10 or more genome-wide significant loci. We found 19 GWA studies and examined whether the studies considered sex differences, and if yes, whether they found significant sex differences.

Author	**Year of publication**	**Phenotypes**	**Were sex differences explored?**	**Sex-specific result?**	**Conclusion: Diff genes M/F?**
Y.S. Cho *et al*	2009	Height, BMI, pulse rate, systolic BP, diastolic BP, waist-hip ratio, bone mineral density	Yes, in 11 top-hits SNPs in stage 1 GWAS and in stage 2 replication analyses	1 SNP (associated with SBP) showed nominal statistical significance, but this sex-specific effect was not replicated. No significant sex-heterogeneity at the height associated signals.	No
A. N'Diaye *et al*	2011	Height (meta-analyses)	Yes, for 2 SNPs	No significant evidence of sex heterogeneity (P=.26 & P=.34, respectively).	No
C.L. Carty *et al*	2012	Height	No (replication analyses of 169 SNPs was carried out using sex- (and disease status-) specific height Z-scores)		NA
D.C. Croteau-Chonka *et al*	2010	Height, weight, BMI, waist circumference	No (the sample consisted of women only).		NA
J.Z. Liu *et al.*	2010	Height, BMI	No (sex and sex by SNPs interaction as covariates)		NA
A. Tonjes et al	2009	Height	1 significant SNP was further considered in two independent cohorts	SNP was replicated in males (p=.049 & p=.022) not in females (p=.965 & p=.055). But no sex differences in the discovery sample	Inconclusive
A. Johansson *et al*	2008	Height	Linkage and GWA were run stratified for sex in 5 European samples and in combined sample	No sex-specific results in linkage or GWA. However, signal strength of the significant SNPs was different in males and females.	No
DF. Gudjartsson *et al*	2008	Height	Yes, effect sizes for the identified SNPs were compared for males and females	No significant differences between sexes (p-values>.002).	No

G. Lettre *et al*	2008	Height	Yes, effect sizes of 16 identified SNPs (combined P < $5X10^{-6}$) were compared for males and females in a replication sample.	No evidence for sex heterogeneity (P > .05).	No
S. Sanna *et al*	2008	Height	2 SNPs were considered	No evidence for sex heterogeneity	No
M.N. Weedon *et al.*	2007	Heigth	1 SNP was considered	No evidence for sex heterogeneity	No
M.N. Weedon *et al.*	2008	Height	Yes, effect sizes of 20 SNPs were compared for males and females in stage 2 joint analyses.	Of the 20 significant loci, 1 SNP was found to have a significant greater effect in females than males (P=.01).	1 SNP larger effect size in females.
H. Lango Allen	2010	Height	Yes, sex-specific analyses of the 180 associated signals in the stage1 + stage2 samples	No differences between their effects in males compared to females (p-values>0.01)	No
HN Kim	2010				
Y. Okada	2010	Height	No (sex as covariate)		NA
J.J. Kim	2009	Height (adults), idiopatic short stature (children)	No (sex as covariate)		NA
K Estrada	2009	Height	No (sex as covariate)		NA
N Soranzo	2009	Height, trunk length, hip axis length and femur length	Yes, sex-specific effects at the 17 validated height loci were investigated in 1 sample (the Rotterdam sample)	Significant sex-heterogeneity at 3 SNPs (p=.01, I^2=85%; p=.002, I^2=89% & p=.002. I^2=89%, respectively);	Significant sex-heterogeneity at 3 SNPs (Nfemales=3,374; Nmales =2,362)
SF Lei	2008	Height	No (sex as covariate)		NA