

## INTERNSHIPS BIOLOGICAL PSYCHOLOGY VRIJE UNIVERSITEIT AMSTERDAM

Please find below the internship positions at the Department of Biological Psychology. If you are interested in doing an internship at this department, please send an email with a motivation letter, CV, and a list of your grades to the Internship coordinator, Dennis van 't Ent (d.vant.ent@vu.nl). If you are interested in multiple topics, please provide a top-3.

### **Investigating the relationship between affective states and autonomic nervous system activity**

In a recently started project in collaboration with Philips, the relationship between (positive & negative) affect and various autonomic nervous system (ANS) activity measures is investigated. The relationship between affect and ANS activity has already been a topic of interest for many years. However, to date the vast majority of these studies have investigated this relationship in a laboratory setting with use of artificial stressors, limiting their ecological validity. This is partially due to a lack of possibilities to measure ANS activity in an ambulatory setting in the past. However, progress is being made in the development of such devices, like the VU-AMS here at the VU. The aim of the current study is dual. Firstly, this study aims to investigate the relationship between affect and ANS activity during daily life with use of the VU-AMS. Secondly, this study aims to validate two wrist watches (one measuring heart rate and one measuring skin conductance) under development at Philips. These watches could provide novel ways to measure ANS activity during prolonged recordings in daily life.

To study these aims participants will perform a variety of task (known to influence ANS activity) at our laboratory while wearing the VU-AMS and the two Philips watches. They will continue to wear these devices during a 24 hour ambulatory measurement to measure ANS activity in daily life. At the laboratory section participants are asked to report their affect after various tasks, by filling in a short questionnaire. During the ambulatory section participants are asked to report their affect (but also main activity, location and social environment) by filling in a short diary on an iPod hourly.

#### **Level: B-thesis & M-thesis**

Students gain hands on experience in collecting and handling diverse data types. The bachelor student will assist the PhD in participant recruitment, data collection and data cleaning and perform analyses on the laboratory data for a B-thesis. The master student will also be actively engaged in data collection but additionally perform more involved data analyses combining data from various domains and across laboratory and ambulatory data sets.

**Requirements:** For the both the B-thesis and M-thesis student this project requires basic statistical skills, knowledge of SPSS and background knowledge on the biology of the ANS. For the master student experience with data collection, VU-AMS or diary data, R/Matlab and fluency in Dutch is preferred.

**Supervisors:** prof dr Eco de Geus, dr. Martin Gevonden & Denise van der Mee, MSc.

### **What is the role of the autonomic nervous system in cognitive functioning?**

There is a growing body of empirical evidence demonstrating that (intra- and inter-) individual differences in cognitive functioning relate to differences in the activity of the autonomic nervous system (ANS). Indices of heart rate variability, for instance, have been found to associate with verbal memory in middle-aged men and with executive functioning in the elderly. In addition, autonomic nervous system activity is associated with the performance on tasks requiring attention and inhibitory control in both healthy and clinical samples of children. The master thesis student will engage in an in-depth investigation of the association between cognitive task performance (speed+accuracy) and cardiac ANS activity at rest and during the actual execution of the cognitive tasks. For this investigation we have already collected a large dataset on cognitive as well as cardiac ANS functioning in a sample of ~1000 twins and twin family members (Swagerman et al. 2015, Brain Cognition, 97:32-9). Measures of cognitive functioning span a broad range of cognitive domains (17 different tasks). Because data were collected in a genetic informative sample they provide the unique opportunity to study genetic and non-genetic contribution to the association. Various student-specific questions on modifiers of the association between ANS and cognition can be formulated, because participants had been interviewed extensively on e.g., education of parents and self, subjective health or well-being, smoking and drinking behavior, exercise behavior, and medicine use.

The student will be tasked with (1) careful signal analysis of the ECG and impedance cardiography

recordings at rest and during the cognitive tasks to extract measures of cardiac ANS activity, (2) performing analyses to confirm or refute an association between this cardiac ANS activity and the performance in various cognitive domains, and (3) to report on the results of the analyses in a master thesis.

**Level: M-thesis**

**Requirements:** Good statistical skills, knowledge of SPSS, and ideally R/Matlab ; background knowledge on the biology of the ANS.

**Supervisor:** Prof dr Eco de Geus

### **The Happiness-Formula**

Recently it has been claimed that there is some kind of happiness-formula (see for example Mo Gawdat's book Solve for happiness). In the presentation of this formula differences between people are largely ignored. Is there such thing as a happiness-formula? And which factor play a significant role. In a collaborative project of Hope XXL ([hope-xxl.com](http://hope-xxl.com)), Professor Meike Bartels, and Bart Baselmans, information in happiness, flourishing, and satisfaction with life is collected in a sample of about 2000 individuals living in the east of the Netherlands. Besides these well-being measures, we gather information on life-style, living environment, and work. In this project, we would like to sort out if some factors are more important than others factors for overall well-being. We, furthermore, would like to compare these data to population-based data of the Netherlands Twin Register. HOPE XXL wants to ensure that all people can achieve a life they can grade as 'good'. Therefore, they developed a new vision on the future with young people from all over the world. This vision, called the Liemers List, was presented to the United Nations in February 2015. With the current project, we will start to make this vision evidence-based.

**Level: B-Thesis**

**Requirements:** basic statistical skills

**Supervisor:** Prof. dr. Meike Bartels & Bart Baselmans.

### **Why do regular exercisers have a higher wellbeing?**

In the adult population regular exercise has consistently been associated with higher levels of wellbeing such as fewer depressive and anxious symptoms and higher levels of subjective wellbeing. Within this thesis project we will add an extra dimension to the research on exercise behavior and wellbeing by including information on the self-perceived exercise ability relative to peers. The hypothesis is that regular exercisers perceive themselves to be better at exercise than non-exercisers, and that the self-esteem derived from this self-perceived exercise ability is a main source of the increase in their subjective wellbeing.

Large survey data on subjective wellbeing, self-perceived exercise ability and regular exercise behavior are available and can be used by the B-thesis student.

With this dataset the following research question can be investigated by so-called mediation analysis: Will the association between exercise and wellbeing disappear if we control for self-perceived exercise ability?

**Level: B-thesis**

**Requirements:** Some experience with basic statistics and SPSS (or comparable programs) is required.

**Supervisors:** Prof dr Meike Bartels & Prof dr Eco de Geus

### **Heritability of objectively assessed physical activity during the week and the weekend.**

Voluntary exercise behavior has been found to be heritable from adolescence onwards. Genetic differences between individuals account for up to 70% of the variance in exercise behavior. These large-scale population based studies though, are based on self-report and survey-based measures of voluntary exercise behavior. A disadvantage with this survey method is that a lot of physical activity is not captured because it occurs outside exercise activities (dancing, biking, physical labor at work). Within this project the student will use data on physical activity collected with a hip-worn accelerometer. Resemblance between MZ and DZ twins will be used to estimate the heritability of objectively measured vigorous intensity physical activity, moderate intensity activity, and light intensity physical activity, separately for week days and the weekend.

**Level: B-thesis**

**Requirements:** Some experience with basic statistics and SPSS (or comparable programs) is

required.

**Supervisors:** Prof dr Eco de Geus & Matthijs van der Zee (PhD candidate)

#### **Exercise behavior and epigenetics in monozygotic twins**

Epigenetic mechanisms such as DNA methylation regulate the expression of genes in cells, and may respond to environmental influences. Previous studies have reported that exercise may induce epigenetic changes in cells. For example, directly after exercise, skeletal muscle cells show changes in DNA methylation and gene expression that trigger structural and metabolic adaptations of muscle tissue. Vice versa, epigenetic mechanisms may also influence exercise behavior. This project aims to investigate if monozygotic twins who differ in voluntary exercise behavior display differences in DNA methylation in white blood cells. This study may provide novel insights into the effects of exercise behavior on immune-system related functions. The student will perform statistical analysis of existing DNA methylation data from monozygotic twins and link this to information on exercise behavior to examine the question which locations in the genome show differences in DNA methylation between monozygotic twins who are discordant for exercise behavior.

**Level: B-thesis or M-thesis.** For a B-thesis, the student will perform basic statistical analyses (paired t-tests, correlations) of candidate genes in R. For an M-thesis, the student will analyze a genome-wide DNA methylation dataset.

**Requirements:** An interest in the topic, experience with basic statistics and some experience with a statistical program such as SPSS or R.

**Supervisor:** Matthijs van der Zee Msc, Dr. Jenny van Dongen, Prof. Eco de Geus

#### **Do taller women have more children? Unravelling the connection between female fertility and body composition using genetic information.**

Previous studies have showed that spontaneous dizygotic twinning is associated with body composition (height, body mass index) and smoking. Mothers of DZ twins differ from other women in the population namely, they are taller, have increased body mass index and are more likely to smoke before pregnancy. However, these results are based on observational studies, and further investigation is needed to disentangle whether these associations are causal in nature. Do taller women have more children? Is there a common genetic cause affecting both body composition and fertility? This study will take advantage of the recent progress made in identifying genes that are related to fertility, spontaneous dizygotic twinning, smoking, and anthropometric traits. The aim of this project is to test these hypotheses using genetic and epidemiologic data on fertility and reproductive behavior collected at the Netherlands Twin Register, as well as summary statistics from large genome-wide association studies of fertility and anthropometric traits from the Twinning Consortium, the Human Reproductive Behavior Consortium, and from the GIANT consortium .

**Level: B or M thesis**

**Requirements:** experience/motivation to learn R programming, SPSS, interest in twin modeling.

**Supervisor:** Dr. Hamdi Mbarek & Dr. Camelia Minica

#### **Epigenetic footprints of smoking: Does smoking cause changes in DNA-methylation and do these epigenetic changes contribute to the development of addiction?**

Epigenetic mechanisms such as DNA methylation regulate the expression of genes in cells, and may respond to environmental influences. Epigenetic mechanisms have been hypothesized to be involved in mediating the addictive effects of cigarettes and other substances by changing long-term gene expression. Previous studies have reported that substance use in humans, such as smoking and alcohol consumption is associated with widespread differences in DNA methylation at many genes. However, it is largely unknown which of these methylation differences have been induced by substance use and which differences contribute to substance use and addiction disorders. This project aims to investigate which loci in the human genome show evidence for a causal effect of substance use (smoking) on DNA methylation, and which loci show evidence for a causal effect of DNA methylation on substance use. The student will test these hypotheses by using a unique novel analysis method that utilizes information on substance use from monozygotic and dizygotic twins, genotype data and epigenetic data (DNA methylation). The student will analyze existing data from the Netherlands Twin Register and publicly available summarized data on genetic associations with smoking and DNA methylation.

**Level: M-thesis**

**Requirements:** An interest in the biological effects of substance use, experience with or motivation to learn programming in R, interest in causality modeling.

**Supervisor:** Dr. Camelia Minica, Dr. Jenny van Dongen

**Why are behavioral control and psychopathology related in childhood?**

Children vary a lot in their ability to adopt to rules and control their emotions, behavior and impulses. This so called behavioral control is a moderate to highly heritable trait with estimates for the heritability ranging from 50 to 80% and shows genetic overlap with psychopathology in adulthood. Much less is known on this link between behavioral control and different types of psychopathology in children. The large dataset available in the Netherlands Twin Register on behavioral control and psychopathology in childhood will make it possible to confirm whether behavioral control and psychopathology, e.g. attention problems, autistic traits, anxiety, depressive symptoms, and what underlies this co-occurrence. Twin studies will be used to determine the contribution of genetic and environmental factors to the association between behavioral control and psychopathology. The classical twin design utilizes the fact that monozygotic (MZ) and dizygotic (DZ) twins resemble each other to a different extent.

**Level: B-thesis**

**Requirements:** An interest in the topic and some experience with basic statistics and SPSS (or a comparable program) is required.

**Supervisor:** Dr Eveline de Zeeuw

**What is the influence of parental psychopathology on the behavioral control of children?**

Parental psychopathology has been shown to be associated with internalizing and externalizing problems in children. However, most research cannot draw conclusions about the causal pathways underlying these associations. It is possible that the association is merely due to the fact that parents do not only provide the environment in which their children grow up, but also each contribute 50 per cent of their genes. When there is an overlap in genes involved in parental characteristics and a child's behavioral control level, we cannot say whether there is a true effect of parental psychopathology on behavioral control unless we take into account this genetic overlap. In this project you will use the children-of-twins (CoT) design which is based on the differences in genetic relatedness between different family members to test the direct effect of parental psychopathology on behavioral control. Children receive 50% of their DNA from either parent which means that siblings with the same biological parents share approximately 50% of their genetic material, while half siblings share 25% and full cousins share only 12,5%. Just like regular siblings, dizygotic (DZ) twins share approximately 50% of their genes, but monozygotic (MZ) twins share almost all their genetic variance. As a consequence, the offspring of MZ twins are as genetically related to the co-twin of their parent (uncle/aunt) as they are to their own parent. However, they only share a home environment with their parents, making it possible to make a distinction between genetic and environmental transmission from parents to their offspring.

**Level: M-thesis**

**Requirements:** An interest in the topic and some experience with basic statistics and SPSS (or a comparable program) is required.

**Supervisor:** Dr Eveline de Zeeuw

**Autism Spectrum Disorder and its comorbidities**

Autism spectrum disorders (ASD) are characterized, in varying degrees, by difficulties in social interaction, verbal and nonverbal communication and repetitive behaviors. ASD has been shown to be associated, amongst others, with intellectual disability, inattention, internalizing problems and difficulties in motor coordination. Polygenic scores are a summary of a person's genetic predisposition for a certain trait based on weighted associations between all, not only significant, genetic variants and a trait. Polygenic scores allow for the exploration of the underlying etiology of comorbidities. Here, it is expected that genetic variants associated with one trait, will explain part of the variance in the other trait. Polygenic score analyses will be used to look at the pattern of comorbidity of ASD and other traits at the measurement genotype level.

**Level: B or M-thesis**

**Requirements:** An interest in the topic and some experience with basic statistics and SPSS (or a comparable program) is required.

**Supervisor:** Dr. Eveline de Zeeuw

#### **Associations between subjective happiness and cortex structure**

Differences in wellbeing between individuals spark reason to study the (neuro)biological foundations in order to gain understanding of the mechanisms that underlie a happy and satisfactory life.

Interestingly, some recent studies pointed to an association of subjective happiness with volumes of parietal lobe structures, especially of the precuneus. In the present study we further investigated possible relations of subjective happiness with regions of the cortex.

**Level: B-thesis**

**Requirements:** Some experience with basic statistics and SPSS

**Supervisor:** Dr Dennis van 't Ent

#### **Depression and social support**

We have found that with a similar genetic risk and with a similar level of environmental risk factors, such as serious life events, people still differ in their risk of developing depressive symptoms. One explanation, initially observed in a small study of identical twins who were discordant for depression, is that they differ in the degree in which they experience social support. In this project, we will look at the causes of individual differences in the extent to which people experience social support, support from their religion/church and the extent to which these are related to or explain differences in depressive symptoms.

**Level: B or M-thesis:** if the thesis is to qualify as an MA project, an additional component can involve working with polygenic risk scores.

**Requirements:** this project requires innovative analysis of existing data, knowledge of SPSS (or similar software) is required, willingness to consider working with clustered (family level) data.

**Research participation:** This project also offers the option to be involved in a national project collecting data on depression in the Dutch population.

**Supervisor(s):** Prof dr Dorret Boomsma

#### **Screenviewing and Childhood and Adolescent Academic Achievement**

TV viewing and use of iPad and computers has steeply increased over the past decade, also in young children. Initiatives to change schools into "Steve Jobs schools" by providing iPads to all pupils, is currently a reality. This change has been mostly interpreted as positive, but this is questionable. Is screenviewing behavior in children a predictor for later academic achievement?

**Level: B or M-thesis:** if the thesis is to qualify as an MA project, additional questions involve bivariate genetic analyses of twin data.

**Requirements:** this project requires innovative analysis of existing data, knowledge of SPSS (or similar software) is required, willingness to consider working with clustered (family level) data.

**Research participation:** an active role in acquiring data from teachers of children

**Supervisor(s):** Prof dr Dorret Boomsma

#### **Family conflict & self-control in adolescence**

Family conflict brings tremendous costs to individuals and society. Although it is reliably associated with multiple problems, the pathways underlying these associations are poorly understood hampering effective interventions. We propose that self-control plays a key role in family conflict because of its foundational function in regulating behavior, emotions, and cognition. As a result, we hypothesize that coping with family conflict and associated stressors depletes individuals' self-control strength. In order to test this hypothesis, this study aims to investigate the association between family conflict and associated stressors and self-control, expecting them to correlate cross-sectionally and longitudinally. The Netherlands Twin Register includes measures to validly operationalize self-control (ASEBA-SCS) family conflict (e.g. "we fight a lot at home", "Family members get so angry, they throw things") and family stressors (e.g. "someone close passed away", "one of my parents lost their jobs recently"). These measures have been assessed at age 14, and 16 and therefore the present study will focus on adolescence.

**Level: B or M-thesis**

**Supervisors:** Yyouk Willems, Msc & Prof dr Meike Bartels

**Reference:**

Finkenauer, C., Buyukcan-Tetik, A., Baumeister, R., Schoemaker, K., Bartels, M., & Vohs, K. D. (2015). Out of control: Identifying the role of self-control strength in family violence. *Current Directions in Psychological Science*, 24, 261-266

**Are optimistic individuals happier or does a half empty glass not influence overall wellbeing?**

It sounds obvious that optimistic people are happier. Or could we better say that happy people are more optimistic? This is however more based on folk wisdom and on any evidence-based fact. Individual differences in optimism and wellbeing, such as happiness, are accounted for by genetic differences and differences in environmental influences. About 40% of the variance is accounted for by genes. The open research questions are if there is a significant association between optimism and wellbeing and if so if this is accounted for by overlapping genes. In addition the large-scale longitudinal data of the NTR provide the opportunity to gain insight into the direction of causation. Do optimistic people become happier or are happy people more optimistic?

**Level: B-thesis**

**Requirements:** Some experience with basic statistics and SPSS (or comparable programs) is required.

**Supervisor:** Prof dr Meike Bartels

**Choosing your environment; Differences in environment between happy and unhappy individuals**

There has been a fair amount of research, including my own work, into the underlying sources of familial aggregation of wellbeing. A recent meta-analysis reveals that about 35% of the variance in wellbeing is accounted for by genes. Additionally, several studies have attempted to unravel the environmental and social correlates of wellbeing. Main focus has been on income and life events, such as relationship status and employment status. A literature review, however, calls for caution in the interpretation of such determinant studies, due to lack of evidence on the direction of causation. It should also be noted that most of these 'environmental' or 'non-genetic' correlates are actually influenced by genes, necessitating an integrated approach of genetic and environmental research. Focus so far has been on additive models in which independent effects of genes and environment add up to explain individual differences in wellbeing. If, however, genes and environment do not act in an additive manner, as is to be expected for multi-factorial complex traits such as wellbeing, ignoring the presence of gene-environment interplay will lead to biased estimates of the relative importance of both genetic and environmental factors and will provide an incomplete and simplistic picture of the etiology of wellbeing. The proposed thesis focuses on the gene-environment correlation in wellbeing. Gene-environment correlation (rGE) describes the genetic control of exposure to the environment, resulting in a non-random distribution of environments over distinct genotypes. Within this project the results of a genome-wide association study for Wellbeing will be used to compose poly-genic risk scores. This risk scores will then be associated to environmental factor as assessed in the large-scale database of the Netherlands Twin Register. With this project we will gain insight into the "happy environment".

**Level: M-thesis**

**Requirements:** this project requires innovative analysis of existing data, knowledge of SPSS (or similar software) is required,

**Supervisor:** Prof dr Meike Bartels & Bart Baselmans, MSc

**Optimizing efficiency of genome-wide association studies for multiple phenotypes** Genome-wide association studies (GWAS) aim at identifying associations between the genome and a phenotype (or trait) of choice, for example having a disease or not. The association is tested between the traits of interest and the typically ~4M to ~12M genetic variants that potentially underlie the observed variation in a trait. A recent development in the GWAS field is using metabolomics data as input phenotypes.

The metabolic traits are based on metabolomics analysis of human blood serum using samples from the Netherlands Twin Register (NTR). The aim of metabolomics is to obtain a comprehensive qualitative and quantitative overview of the substrates, intermediate and end products of cellular metabolism as can be measured in, for example, body fluids. The number of metabolites (for example, lipids, sugars) detected in a single sample from a single individual using metabolomics can easily run into the order of hundreds to thousands.

Because of the massive numbers of association tests, parallel computing on high-performance clusters

is often warranted, even when performing GWAS for only one or a few traits. Adding to this, it has been shown that ~50% to 85% of all statistically significant hits in GWA studies for metabolomics reside in the ratios of the detected single metabolites. These ratios are a proxy for enzyme activity under steady-state conditions. For determining the genome wide associations for all  $N$  single metabolites *and* their ratios, the total number of metabolic traits to be tested in the study is  $N+N(N-1)/2$ . In previous efforts by the NTR, where  $N = 163$  single metabolites were detected in the blood serum sample of each participant, we conducted GWAS of  $N+N(N-1)/2 = 13,366$  metabolic traits. We addressed this computational challenge by adopting a recently developed method for fast GWA: FaST-LMM. In addition, we applied a massively parallel approach on the national compute cluster Lisa. We ran these associations within approximately three days of wall-clock time. This number of computations would have taken one month using conventional GWAS tools. This approach enables the genome wide analysis of the associations between all included genetic variants and all included single metabolites and their ratios in metabolomics-based GWAS studies.

Still, there is room for improvement. FaST-LMM performs an analysis for each trait separately. This means that for each of the 13,366 traits the genotype data has to be read into memory. Genotype data files are typically large, hence it would be far more efficient to read in these data once and to perform the associations iteratively over the input traits. The FaST-LMM source code (available online) provides a framework for handling multiple traits. However, this has not been implemented. The goal of this project is to implement a 'multiple trait' option in FaST-LMM. If this can be achieved, a large scale GWAS will be conducted.

The work plan consists of Dry-lab computational work mainly involving programming and testing/benchmarking. The work is of practical nature. However, if sufficient progress is made, GWA analyses can be performed on a large scale. The latter makes the work scientifically more interesting for Master and Bachelor students.

**Level: B or M-thesis**

**Requirements:** preferably following study in HBO–bioinformatics/HBO–computer science. Skills required:

Programming in C/C++ (Preferably an object-oriented programming language). Programming can be performed in MS Visual Studio environment, although the Linux platform (gcc/g++) is preferred. Experience with Linux and/or cluster computing. Experience in data analysis.

**Supervisor:** Dr René Pool

**Are substance users more likely to develop mental illness?**

Substance use (cannabis, smoking, alcohol) is increasingly recognized as a major, serious public health concern. Strong evidence suggests that early initiation of substance use (before age 18) '*opens the gate*' for experimentation with other drugs and progression to escalated use, abuse and dependence. Regular substance use is predictive of health problems, poor educational and professional achievement. Recent evidence suggest that regular substance users are more likely to develop mental illness (e.g., psychotic, depressive) symptomatology relative to non-users.

However, what we know about the relationship between substance use and mental disorders is mainly based on observational evidence. A major problem is that this type of evidence says nothing about whether the association is causal in nature, or about direction of causation (e.g., does mental disorder symptoms cause increase the likelihood of initiation/regular substance use, or does substance use causally affect the risk of developing mental illness symptoms?). This project aims to disentangle the nature of the association between substance use and mental illness symptoms (is the association causal in nature?), and to test whether mental illness symptoms are causes or consequences of substance use. The student will test these hypotheses using a recently developed and very popular analysis method that exploits genetic information to test causal hypotheses. The student will analyze data from the Netherlands Twin Register and publicly available summarized data on genetic associations with substance use and mental illness symptoms.

**Level: B/M-thesis**

**Requirements:** A motivated student with high interest in the mental health effects of substance use, experience/motivation to learn R programming, interest in twin modeling.

**Supervisor:** Dr. Camelia Minica

**Effect of genetic principal components in twin modeling**

Within genetic research to find genes it is common practice to add principal components based on

SNP data as covariates, in order to correct for possible population stratification. Population stratification can be present at all levels within the studied population, either based on different ethnicity, but also based on other factors like religion or geographic location. Recently a whole new wave of heritability analyses has arisen, which tries to estimate all the genetic variance explained by the measured common SNPs, the GCTA approach. Results from this heritability modeling has shown very different outcomes. An interesting part of these analyses is, that here often PCs are taken into account as well. This in contrast to the classic twin heritability models where this is not the case.

Within the Netherland Twin register we have data on multiple phenotypes, we have GWAS SNP data and we have twin families. With the help of an empirical example and simulations of various conditions we want to see the effects on twin heritability, when we are correcting for PCs like GWAS and GCTA. This can be a Bachelor or Master thesis subject. The student will learn how to perform various heritability analyses and to make simulations of the various studied models.

**Level: B or M-thesis**

**Requirements: -**

**Supervisor:** Prof Conor Dolan & Dr Jouke- Jan Hottenga

#### **Effects of genome wide SNP imputation on the ancestry estimates of individuals**

In genome wide association studies, genotypes are imputed to allow for subsequent meta-analysis of the data, and to cover the whole genome with all viable known variants. In addition, principal components (PCs) are often calculated from the SNP data of individuals to estimate their ancestry. These PCs are then used in the association analysis to correct for population stratification. The imputation of genotypes is usually based on a reference panel with mixed ancestry as studies have shown that the results are optimal in this case. An interesting question is however, if the ancestry estimates are also affected by the imputation itself. A further question is how serious an imputation is affected when the ancestry is absent in the reference panel, and another question related to this is why there are small differences between ancestry estimates in monozygotic twins. In order to answer these questions sequence and chip platform data from the Netherlands Twin Register will be (re-)imputed and the effects on the PCs will be calculated. Furthermore we could for example impute the Neanderthal genome and see the ancestry estimates of this person. The methods learned in this research are cleaning and analyzing genetic SNP & sequence data, determining ancestry by PCs and improve scripting and computer skills.

**Level: B or M-thesis**

**Requirements: -**

**Supervisor:** Dr Jouke- Jan Hottenga.

#### **Batch effect in the classical twin design.**

Biological phenotyping often involves the processing of biological samples in batches. Batch processing may give rise to batch effects. In downstream statistical analyses, it is important to accommodate batch effect, as they are a source of data clustering. In the classical twin design, such clustering may manifest as false shared environmental effects. Batch-based phenotyping in the classical twin design poses a design issue concerning the allocation of twins to batches. Specifically is it better to allocate twin pairs to batches or individual twins to batches?

**Level: B-thesis**

**Requirements:** Interest in statistical modeling of twin data, experience with the R programming environment

**Supervisors:** Prof. Conor Dolan

#### **Using the liability threshold model to address scale dependency in tests of GxE interaction in the moderated classical twin design.**

The classical twin design has been extended to include moderations effects, where the genetic and environmental effects of one phenotype (Ph) are a function of a measured moderator (M). This approach to the detection of moderation is known to be scale dependent. That means that the detect moderation may is a function of the (non-normal) distribution of the data, and not of true moderation. The liability threshold model does not rely on the observed distribution of the data, and therefore may provide a useful model to investigate moderation. Is this a viable approach to semi-scale independent moderation analysis?



**Level: B-thesis**

**Requirements:** Interest in statistical modeling of twin data, experience with the R programming environment

**Supervisors:** Prof. Conor Dolan, Dr. Michel Nivard

**Within person Phenotype to E transmission difficult... within person A to E transmission doable?**

Phenotype to E transmission in the repeated measures twin design amounts to the transmission of Phenotypic effects at time t on the environmental variable at time t+1. Specified as a within twin phenomenon, this is hard to detect, due to lack of information. How does A to E transmission fare, where there is a direct relationship between the additive genetic variable at time t and the environmental variable at time t+1?

**Level: M-thesis**

**Requirements:** Interest in statistical modeling of twin data, experience with the R programming environment (and the OpenMx library)

**Supervisor:** Prof. Conor Dolan

**Can we relate the cell specific transcriptome of brain cells to specific psychiatric and neurological disorders?**

In this project you use 2 large data resources: 1. Public genome wide association result for 12 psychiatric and neurological disorders 2. The RNA expression profiles of individuals brain cells (e.g. Neuron, Microglia, Astrocytes etc. etc.), obtained from either fetal of adult tissue and from 2 brain regions From this last resource we can determine which genes are specifically overexpressed in certain cell types, during a certain phase in live, or in a specific brain region. Subsequently we test where sets of genes associated with a specific cell type are differentially enriched in their effect on psychiatric and neurological disorders.

**Level: M-Thesis, Master-internship, B-Thesis**

**Requirements:** t.b.c.

**Supervisor:** Dr Michel Nivard

**Genetic factor scores in the genomic age**

GCTA is a linear mixed model used to estimate the variance attributable to a large number of measured genetic variants. Traditionally such linear mixed modeling was used for prediction in animal breeding, where the aim was to determine the breeding value of the animal or plant. In GCTA, as applied to human data, this aspect is not considered, as the focus is on the estimation of (chip-based or genomic) additive genetic variance. The aims of this project is 1) to consider the estimation of breeding values following GCTA analysis in a manner analogous to that used in animal breeding studies, and 2) determine the relationship of these with additive genetic factor scores as calculated in the classical twin design, 3) to determine the use of breeding values in studying genotype environment interaction and covariance.

**Level: B or M-thesis**

**Requirements:** interest in structural equation modeling and linear mixed modeling

**Supervisor(s):** Prof Conor Dolan & Dr Jouke-Jan Hottenga

**Heritability of motor milestones**

Motor milestones are recognizable landmarks of development in the first years of life. A delay in achieving motor milestones is often a first sign for further monitoring the development of a child. One of the risk factors for a delayed motor development is prematurity, especially very preterm birth is associated with impaired motor development (de Kieviet et al. 2009). The goal of this study is to estimate the relative contribution of genetic and environmental factors on achievement of motor milestones and to test whether these factors contribute differently in term, preterm and very preterm born twins.

The study is part of the longitudinal data collection of the Young-Netherlands Twin Registry on behavioral and emotional development. Data on motor development were available for about 21,500 twin pairs, born between 1987 and 2010, and obtained by maternal report.

**Level: B-thesis**

**Requirements:** Some experience with basic statistics and SPSS (or comparable programs) is required.

**Supervisors:** Dr. Toos van Beijsterveldt

### **Are we programmed before birth to respond more or less to stress?**

Some people respond strongly to stress, others remain physiologically calm. It has been suggested that our stress response may be determined by our experiences in the womb. This would result in an association of the stress response with birthweight and gestational age. One way of studying this association is to examine the difference within twin pairs, who share many of the same experiences, but can differ in birth weight. Within the Netherlands Twin Register twins and family members (N~1400) underwent 24-hour ambulatory monitoring of heart action and provided cortisol samples across the day. For most of these individuals information on gestational age and birthweight is also available. You will explore the association between prenatal circumstances, as indexed by birth weight and gestational age, with physiological stress profile, making full use of the fact that with the inclusion of (monozygotic) twin pairs control for genetic and environmental factors is possible.

**Level: B or M-thesis**

**Requirements:** This project requires statistical skills and knowledge of SPSS (or similar software).

**Supervisors:** Dr Gonneke Willemsen

### **Are twins really the same as singletons?**

When you ask behavioural geneticist, they will state that twins are not different from singletons. This assumption is one of the foundations for the study of heritability using the twin design. However, when you ask twin families some will state that twins are different, an opinion that may be shared by the general public. The Netherlands Twin Register offers the unique possibility to explore this assumption for a wide range of factors in the domain of development, lifestyle, psychology, physiological and disease. Depending on your interests and level of study you will focus on one or more of these domains, exploring the literature and testing the differences between twins and their singleton siblings.

**Level: B or M-thesis**

**Requirements:** This project requires statistical skills and knowledge of SPSS (or similar software).

**Supervisors:** Dr Gonneke Willemsen & Prof. dr. Dorret Boomsma

### **Where we live may be just as important as how we live.**

The linkage of geographical information to databases enables us to study environment far more extensively than in the past. Within the Netherlands Twin Register we have information on lifestyle and wellbeing but now also on a large number of neighbourhood characteristics. This allows for a myriad of questions to be answered. For instance, you may examine the interaction between lifestyle factors such as exercise behaviour and smoking and neighbourhood factors such as sport facilities and pollution levels on various health indicators. After an exploration of the literature and the available geographical and health indices you will decide upon the topic you want to focus on. In this you will make full use of the possibilities that the inclusion of (monozygotic) twins in your dataset offer you. This may include studying the heritability of the geographical indices or applying the cotwin control design.

**Level: B or M-thesis**

**Requirements:** This project requires statistical skills and knowledge of SPSS (or similar software).

**Supervisors:** Dr Gonneke Willemsen

### **Causes of variation in cholesterol level: genes, food preference and regional differences**

During a large-scale blood collection in the Netherlands Twin Register research nurses observed differences in the blood samples across the Netherlands and hypothesised that the consumption of sea food may underlie differences in cholesterol level. In this study you will determine whether the suggestion of regional differences in cholesterol levels holds when you test this scientifically. In addition, you will explore the interactions between genetic susceptibility for high cholesterol, food preference and regional characteristics. You will combine cholesterol data as obtained in biobanked individuals with information of region of residence, survey data on food preference and polygenetic risk scores. Using regression procedures you will explore their associations, paying particular attention to interactions.

**Level: B or M-thesis**

**Requirements:** This project requires statistical skills and knowledge of SPSS (or similar software).

**Research participation:** scoring of data on employment / profession

**Supervisors:** Dr Gonneke Willemsen

**What makes you decide on the type of job you want? The effects of genes and environment on career choice**

We often choose similar careers as our parents. This may be due to our familiarity with the work our parents do, but it may be also driven by genetic factors. In this study you will use data collected in adult twin family members to study the influence of genes and environment on job choice.

As part of the project you will score job description data collected in the Netherlands Twin Register. These scored data you will use to conduct twin-family analyses. You may also explore assortative mating for career choice, depending on B or M thesis level.

**Level: B or M-thesis**

**Requirements:** This project requires statistical skills, knowledge of SPSS, and knowledge of twin modelling

**Supervisors:** Dr Gonneke Willemsen & Dr Conor Dolan

**Why do some monozygotic twins differ in socioeconomic status?**

Intelligence is one of the most heritable traits and strongly associated with the level of education and occupation that one achieves. Despite this high heritability, there are still twins pairs who differ in their educational or occupational level. Since they both grow up under the same family circumstances, and share the same genes, the question is what factors explain this difference in socioeconomic status? You could think for instance of history of disease (e.g. one of the two could have had an accident) or life events (one of the two had young a relationship or children). Your task is to think of plausible and testable explanations for these differences within monozygotic twin pairs and test these using the large scale data from the Netherlands Twin Register.

**Level: B or M-thesis**

**Requirements:** -

**Supervisor:** Dr Gonneke Willemsen

**Eating behavior and overweight in a Dutch twin family population**

In 2008 Maruyama showed for a Japanese population that overweight was related to two simple questions about eating behavior ("Do you continue eating when you feel full?" and "How quickly do you eat?"). These questions were included in the NTR survey of 2009. The answers of the twins and their family members to these questions can be used to determine whether the same associations are found in a Dutch population, and whether there are age or gender differences. Via genetic behavior studies, you can also determine the heritability of eating behavior, while the relation between the genetic risk profile for BMI with eating behavior can also demonstrate the role of genes for BMI in eating behavior. Other questions that you can think of: is there assortative mating (to what extent are partners alike), are there social factors like educational attainment or living circumstances that may be important?

**Level: B or M-thesis**

**Requirements:** -

**Supervisor:** Dr Gonneke Willemsen

**The effect of work stress and family situation in working women**

As part of an international study, 160 women who worked at the VU and surrounding areas provided saliva samples to determine cortisol and alpha-amylase, during a working day and a weekend day. During these two days, participants also underwent 24-hr ambulatory monitoring during a working day and a weekend day. In addition, they completed a survey about work stress, family situation, depression and lifestyle. These data can be used to determine the effect of work stress and family situation on physiological parameters and depression. Dependent on the nature of your thesis (Bachelor or Master) you will choose one or more aspect to examine. To answer your research question(s) you will need to combine and check the data, before starting the analyses. In the analyses you will also take age, educational attainment and number of working days into account.

NB. For a large number of variables we also have data from London and Budapest. If you are interested, you can also compare the outcomes for the project in Amsterdam with the outcomes in those two cities.

**Level: B or M-thesis**

**Requirements: -**

**Supervisor: Dr Gonneke Willemsen**