Chapter 1.

Introduction
Nicotine is one of the most frequently used addictive substances and, through cigarette smoking, a major contributor to morbidity and mortality. Worldwide, smoking causes 6 million deaths each year (1). While smoking rates have gone down in the Netherlands over the last few decades, still 28% of men and 22% of women were smokers in the year 2014 (2). A better understanding of the causes and consequences of smoking could help to further decrease smoking and thereby improve public health. The aim of this thesis is to explore environmental and genetic influences on addictive behavior with a focus on smoking. It is well known that smoking co-occurs with the use of substances such as alcohol and cannabis (3, 4). Much less clarity exists regarding the relationship between smoking and caffeine use and the relationship between substance use and the consumption of sugar (considered by some as potentially addictive). These two relationships are thus a focus of this thesis. Another key point of this thesis concerns the effect of smoking on mental health. Smoking has been robustly associated with the two most common mental disorders; depressive and anxiety disorders (5, 6). Another disorder that affects more smokers than nonsmokers is ADHD (attention-deficit/hyperactivity disorder) (7). Some evidence suggests that smoking causally increases ADHD symptoms, but this is limited to animal research (8). To further explore this and the other issues described here, data from a unique and large sample of twin families registered at the Netherlands Twin Register (NTR) are utilized.

Smoking behaviour and risk factors

In the Netherlands, 28% of men were current smokers in 2014 while an additional 32% were former smokers. For women, these percentages were slightly lower at 22% and 28%, respectively (2). Male smokers also smoked more cigarettes per day (\(n = 11.4\)) than female smokers (\(n = 10.0\)). When combining the statistics on current and former smoking, the majority of men (60%) and half of women were considered ever smokers, meaning that they smoked regularly at some point in their life-time. The remaining 40% of men and 50% of women were consequently never smokers; those who had never regularly smoked. Smoking rates differed across age groups. Only 5% of the 12-16 year olds smoked in 2014 while this was 23% in the age category 16-20 years. Smoking was most prevalent in 20-30 year olds (37%) and the prevalence was as low as 9% in those aged 75 years or older. When categorizing the Dutch population into four levels of educational attainment going from low to high, smoking prevalence was 30%, 27%, 20% and 16%, respectively (2).

All aspects of smoking behaviour show individual variation and a large body of research has identified risk factors that are associated with smoking initiation, smoking quantity/nicotine dependence, and smoking cessation. Smoking is usually initiated during adolescence, at which age peers are very important and peer pressure to smoke can affect adolescent behaviour (9). In an adolescent sample from the NTR, smoking status of friends was much more predictive of adolescent’s smoking than the smoking status of parents (10). This is probably because adolescents model themselves more to their peers, who are of the same age, than to their parents, who differ from them in age. Low correlations between smoking behaviour of parents
and offspring were also found in an earlier NTR-study. These correlations were not dependent on the sex of the parent or offspring and the resemblance between parents and offspring was explained entirely by genetic relatedness (11). In another, longitudinal sample of Dutch adolescents, a decrease in refusal self-efficacy (the confidence an adolescent has in his/her ability to stay a nonsmoker and refuse a cigarette) predicted smoking initiation (12). Besides the influence of peers, there are individual characteristics associated with an increased risk of taking up smoking. Young males were more likely to initiate smoking than females (13) and a lower education in adolescence was strongly associated with a higher chance of smoking in adulthood (14). Smoking was also more often initiated by Dutch (young) adults living in deprived areas compared to those living in affluent areas, even after correcting for education and income (15). Personality traits are important, with individuals who are more impulsive or prone to experiment being more likely to light up a first cigarette (16) and to do so at a younger age (17). Finally, being a regular smokers was associated with higher levels of extraversion and neuroticism and with lower conscientiousness (self-control and allegiance to social norms) (18).

Once (regular) smoking has been initiated, several factors are associated with individual differences in the number of cigarettes smoked per day and the degree of dependence to smoking. An often used measure of nicotine, or smoking, dependence is the Fagerström Test for Nicotine Dependence (FTND) (19). In a sample of approximately 2,500 current and former smokers registered at the NTR, FTND scores were not associated with age or gender but did show a low, negative correlation with age at first cigarette and a positive correlation with the total number of years a person had smoked (20). In addition, fewer years of education, a lower income and a lower occupation were all associated with higher smoking heaviness and/or FTND scores (21). When comparing light (nondependent) smokers to heavy smokers, the latter reported higher perceived stress than the former while there was no difference between the two groups in level of impulsivity (22).

Most smokers want to quit smoking and have attempted to do so at least once (23). In Dutch current smokers in 2014, 30% of men and 38% of women said to have had a (unsuccessful) quit attempt in the past 12 months (2). Multiple factors are related to the chance that a smoker quits (smoking cessation) or continues to be a smoker (smoking persistence). For instance, individuals who successfully quit smoking were less likely to report symptoms of emotional distress, had a higher self-reported health, drank less alcohol and reported less medical conditions (24). Other predictors of successful smoking cessation were a higher age and a higher educational level while higher FTND scores were associated with smoking persistence (25). Finally, higher levels of neuroticism predicted smoking relapse in former smokers (18).

**Genetic underpinnings of smoking**

Genes play an important role in smoking behaviour and twin studies have been crucial in
estimating how much of the variation in the different aspects of smoking is due to genetic factors. The main premise of the twin model is that the resemblance between two types of twins is compared: monozygotic twins (MZ; share ~100% of their segregating genes and shared environment) and dizygotic twins (DZ; share ~50% of their segregating genes and shared environment). When MZ twins are more similar than DZ twins for a particular trait, genetic influences are implied. When the correlation between DZ twins is larger than half of the correlation between MZ twins, this suggest that there is an influence of the common environment that the twins share. In the NTR it was demonstrated that individual differences in smoking initiation were explained for 44% by genetic factors (26). Most of the remaining variation was explained by common environmental factors shared by the twins (51%), while a very small part was due to unique environmental factors (5%). This moderate influence of genes on whether or not someone starts to smoke is most likely mediated through personality traits such as impulsivity and extraversion which increase the chance of smoking initiation and are moderately to highly heritable (27, 28). Individual differences in smoking heaviness are for the most part genetic in nature. In Dutch twins, 75% of the variation in nicotine dependence was explained by genetic factors with the remaining 25% being due to unique environmental factors (26). There was no influence of the common environment that the twins share. Lastly, approximately half of the individual differences in the ability to quit smoking (smoking cessation) was due to genetic factors in a Finnish twin study, while the other half was due to unique environmental factors (29).

The above described twin studies demonstrate that the phenotype smoking is moderately to largely influenced by a person’s genotype. With the introduction of genome-wide association studies (GWAS), a hypothesis free method to search for genetic variants associated with a complex trait such as smoking became available (30). In GWAS, hundreds of thousands of single nucleotide polymorphisms (SNPs) are measured across the genome. A SNP is a single nucleotide in the genome that is polymorphic, meaning that more than one form is common in the population. For example, at a particular location, or locus, most people have the letter G while a minority of the population has the letter A. In GWAS, the frequency of all included SNPs is compared between individuals with a certain trait or condition (cases) and those without it (controls). In the year 2010, three large GWA meta-analyses were published, investigating the genetics of smoking behaviour (31-33). In a set of pooled analyses of these three studies, several genome-wide significant ‘hits’ were found for smoking behaviour. First, four loci were associated with the number of cigarettes smoked per day. The strongest of these associations was found for rs1051730 which is located in the nicotinic receptor gene CHRNA3. The A allele of this SNP was associated with increased smoking heaviness. Rs1051730 is in very high linkage disequilibrium (LD) with rs16969968 (meaning that these SNPs are usually transmitted together). The CHRNA3 gene codes for the expression of nicotine receptors in the brain, thus providing a plausible explanation for its association with smoking heaviness. For smoking initiation, eight SNPs reached genome-wide significance. The strongest effect was found for rs6265, with carriers of the C allele being at increased risk of smoking.
This SNP is located in the *BDNF* gene, which codes for a neurotrophin that regulates synaptic plasticity and the survival of cholinergic and dopaminergic neurons. It is highly expressed in the prefrontal cortex and hippocampus. These brain regions had previously been found to affect cognitive-enhancing effects of nicotine. Lastly, one SNP was genome-wide significantly associated with smoking cessation (being a former vs. a current smoker). The G allele of SNP rs3025343, located near the *DBH* gene, was associated with an increased odds of successful smoking cessation. The *DBH* gene codes for a protein that converts dopamine into norepinephrine.

The introduction of GWAS also made it possible to estimate how much of the variation in smoking is explained by all of the measured SNPs. Lubke *et al.* (2012) utilized two methods to estimate such SNP-based heritability for smoking, one developed by Yang *et al.* (34) and one by So *et al.* (35). When applying both of these methods Lubke *et al.* (2012) found a heritability of 19% and 28%, respectively for smoking initiation and of 24% and 44%, respectively for current smoking. Corresponding heritability estimates from twin studies were 44% for smoking initiation and 79% for current smoking (36). These findings show that with currently available genotype data, it is possible to explain a considerable part of the heritability of smoking behaviour. However, much of the heritability as found by twin studies remains unexplained. Possible explanations for this so-called ‘missing heritability’ are that twin/family studies have overestimated heritability, that there are many causal variants which each explain a tiny amount of the variation and they therefore do not reach genome-wide significance and/or that causal variants are not in sufficient LD with the SNPs that are genotyped and therefore their effects are not fully captured. More research is needed to uncover the exact explanation (37).

**Smoking and other addictive behaviours**

Several traits co-occur with smoking, meaning that they are present more often in current smokers compared with never smokers (with former smokers often showing intermediate levels). Generally, there are two mechanisms that can explain such an association. 1: A causal effect of smoking on the co-occurring trait or of the co-occurring trait on smoking. 2: The two traits have common genetic or environmental influences. The most prominent association is the one between smoking and the use of other addictive substances, such as alcohol and cannabis. In a large review including 56 studies from around the world the majority reported a strong correlation between alcohol and smoking (4) and in an American study 90% of cannabis users reported that they smoked at some point during their life, compared with 47% of non-cannabis users (3). Less is known about the association between smoking and caffeine use. The strongest contributor to human caffeine consumption is coffee, which showed a heritability of 39% in Dutch twins (38). The influence of genetics on coffee use is thus moderate compared to genetic influences on smoking. This difference in heritability may be due to the fact that caffeine is much less addictive compared with nicotine (39, 40), and heaviness of caffeine use is therefore determined more by environmental factors. Strong
observational associations have been found between smoking and coffee use \((41-43)\). For instance, in an American sample 4.8% of men and 8.1% of women who never drank coffee were smokers, compared with 34.7% and 48.1%, respectively in men and women who drank 6 or more cups of coffee per day \((41)\). Investigations on smoking and caffeinated drinks other than coffee, such as tea, cola and energy drinks, are scarce. Since caffeine is the most used psycho-active substance worldwide \((39)\), a better understanding of the association with smoking is needed.

Besides co-occurring with the use of ‘conventional’ addictive substances, smoking is positively associated with the consumption of sugar \((44)\), a nutrient that is considered by some as potentially addictive \((45)\). Alcohol or drug dependent individuals also have a higher sweet preference than individuals who are not substance dependent \((46-48)\). The consumption of sugar contributes greatly to the rising prevalence of (morbid) obesity worldwide \((49)\). This was for instance shown in a randomized controlled trial where the consumption of sugar through drinks caused weight gain and fat accumulation \((50)\). To date, there is very little research on the association between substance use on the one hand and sugar consumption/liking on the other hand. Interestingly, the consumption of sugar promotes the release of dopamine in the brain and thus has rewarding properties similar to substances such as nicotine or alcohol \((51)\). Given these overlapping effects on the brain’s reward system, sugar consumption and substance use may have common genetic foundations. Twin data are perfectly fit to test mechanisms underlying the association between these two traits.

**Consequences of smoking**

Smoking is a major cause of morbidity and mortality, with some of the most severe consequences being lung cancer \((52)\) and cardiovascular disease \((53)\). Furthermore, smoking has been shown to be correlated with mental health such that smokers are diagnosed with depressive and anxiety disorders more often than nonsmokers \((5, 6)\). Smoking also co-occurs with less prevalent mental disorders such as ADHD. Significantly higher smoking rates have been found in individuals diagnosed with ADHD compared to those without the disorder \((7)\), with one study finding that 40% of adults with ADHD smoke against 26% of the general population \((54)\). It is often assumed that the explanation for this association is that individuals with ADHD or attention problems are more likely to initiate smoking. There is supporting evidence for this explanation from longitudinal studies showing that ADHD leads to smoking, also referred to as the ‘self-medication’ hypothesis \((55, 56)\). Recently, animal research provided compelling evidence for an additional explanation, namely that cigarette smoking causally increases attention problems. In rats, exposure to nicotine during adolescence lead to a decrease in attentional performance, which lasted into adulthood \((8)\). Evidence for such a causal mechanism is not yet available from human studies but could be tested with data of twins.
Content of this thesis
Given the current state of knowledge, there are some unresolved questions. These questions can be addressed by utilizing data of twin-families. To this end I analyzed previously collected data from the NTR and I collected and analyzed new data on addictive behaviour, including smoking, caffeine use and sugar consumption. A brief description of the content of each of the chapters in this thesis is provided below.

Chapter 2 describes the large-scale data collection that has taken place within the Netherlands Twin Register (NTR) as part of this PhD project. Data collection comprised of a survey containing questions on health, personality and behaviour, sent in 2013 and 2014. Approximately 20,000 NTR participants participated in the study by completing the survey. This chapter gives an elaborate description of the methods of data collection and an accurate account of the response rate. In the following chapters of this thesis, previously collected data from the NTR as well as these newly collected data are utilized.

Chapter 3 explores ‘smoking expectancy’, a measure that is obtained by asking people whether they think they will smoke in a year’s time, with answer categories ranging from ‘certainly not’ to ‘absolutely yes’ on a 5-point scale. The meaning of the answer to this question differs depending on whether the person in question is a never smoker (expectancy to initiate smoking), current smoker (expectancy to continue smoking) or former smoker (expectancy to take up smoking again). In a longitudinal design, it is tested whether such a relatively simple question can predict future smoking behaviour. These analyses are corrected for age, gender, educational attainment, self-reported health and smoking quantity and frequency. By employing data of twins, it is also estimated whether a person’s ability to correctly predict future smoking behaviour is influenced by genetic and/or environmental factors.

Chapter 4 describes a study on spousal resemblance for smoking. Spouses resemble each other more than would be expected by chance with the strongest spousal correlations being found for smoking behaviour (57). As of yet, the nature of this association is largely unclear. There are three possible mechanisms that are most often referred to as underlying spousal resemblance. First, spouses may resemble each other due to phenotypic assortment in which case someone’s choice of spouse is directly based on phenotype. Second, social homogamy could pose an explanation, meaning that spouses resemble each other because they are from similar (social) surroundings and were therefore more likely to meet and pair up. Third, spouses may resemble each other because they influence each other while being in a relationship together, in which case there is marital interaction. In this chapter the exact mechanism behind spousal resemblance is elucidated by utilizing data from a large sample of twins, spouses of twins and parents of twins. The effects of research cohort (time of data collection) and age of the participants on spousal resemblance are also explored.
Chapter 1

Chapter 5 gives an extensive account of the observational association between smoking behaviour and caffeine consumption. While a correlation between smoking and coffee use has often been reported, associations between smoking and other caffeinated drinks (tea, cola, energy drinks) are less clear. In addition it is unknown if such associations are consistent across (European) countries with contrasting patterns of caffeine consumption, such as the Netherlands (a ‘coffee drinking’ country) and the United Kingdom (a ‘tea drinking’ country). This chapter provides an answers to these two research questions in two large, population-based samples, one Dutch and one British.

Chapter 6 continues with the topic of smoking and caffeine consumption. There is contradictory evidence on the nature of this association. High observational correlations between smoking and caffeine could be due to causal effects of smoking on caffeine or vice versa, or due to an overlap in genetic and/or environmental factors. This chapter describes a study where three different methods are utilized to test these hypotheses, namely bivariate twin modeling, LD Score regression and Mendelian randomization. With bivariate twin modeling and LD Score regression it can be tested whether genetic and/or environmental risk factors for smoking overlap with genetic and/or environmental risk factors for caffeine use. Mendelian randomization analysis tests whether there are causal effects.

Chapter 7 turns the focus to the association between substance use and the consumption of sugar, a nutrient which some claim has addictive potential (45). In this study, bivariate twin modeling is employed to explore the association between substance use (smoking, alcohol, cannabis, caffeine and illicit drugs) and sugar consumption through drinks. By using data of twins, it can be tested whether genetic and/or environmental risk factors for substance use overlap with genetic and/or environmental risk factors for sugar consumption. This endeavor will further our understanding of the etiology of different types of addictive behaviours, among which the excessive consumption of sugar.

Chapter 8 aims to confirm an important finding from animal research that suggests smoking during adolescence causally increases attention problems (8). It was previously known that smoking and ADHD symptoms, or attention problems, show a high correlation. A commonly posed explanation for this association was that ADHD symptoms causally increase smoking (self-medication hypothesis). Animal research has thus pointed to another option; smoking affects the developing brain and thereby increases attention problems. As of yet there is no such evidence from human studies. This chapter describes the first human, longitudinal study in twins investigating the causal effect of smoking on attention problems, utilizing MZ twins who are discordant for smoking (one twin smokes while the other doesn’t). Twins from these discordant twin pairs are compared on attention problems. When the smoking twin has more attention problems than the non-smoking co-twin, a causal effect of smoking is suggested. Because MZ twins share ~100% of their genetic make-up and a large part of their environment, the design corrects for genetic and most environmental factors.
Chapter 9 concludes this thesis with an overall summary and a general discussion.