RISK FACTORS FOR CHILDHOOD PROBLEM BEHAVIOR:

STUDIES IN TWINS AND TRIPLETS

Diane Lamb

TABLE OF CONTENTS

CHAPTER 1	General introduction………………………………………………………………	9
CHAPTER 2	The heritability of anxious/depressive and withdrawn behavior is under age-related change during adolescence………………………	17
CHAPTER 3	Gene-environment interaction in teacher rated internalizing and externalizing problem behavior in 7 to 12 year old twins............	33
	Supplement to chapter 3..	48
CHAPTER 4	Child care and problem behavior: a study of gene-environment interaction in young Dutch twins………………………………………………	51
CHAPTER 5	Birth weight in a large series of triplets……………………………………	69
CHAPTER 6	Effects of chorionicity and zygosity on triplet birth weight…………	89
CHAPTER 7	Summary and general discussion………………………………………………	109
CHAPTER 8	Nederlandse samenvatting (Dutch summary)……………………………	117
	Appendix I: triplets in the Netherlands Twin Register ……………	127
	Appendix II: Vragenlijst voor moeders van drielingen …………….	132
	Appendix III: Overeenkomst en verschil in uiterlijk van de drieling	144
	Appendix IV: Drielingen: opvoeding, ouderschap en welbevinden	146
	References………………………………………………………………………………	150
	Dankwoord………………………………………………………………………………	167

CHAPTER 1:

GENERAL INTRODUCTION

The nature-nurture debate has developed from a debate questioning whether nature or nurture is influencing a trait to a debate on how nature and nurture together influence a trait, including their additive effects, their interactions and their interdependencies. The objective of this thesis is to make a contribution to this debate in the field of child and adolescent problem behaviors, with a focus on Internalizing (INT) and Externalizing problems (EXT) in children and adolescents. The data on these dimensions of emotional and behavioral problems come from the Netherlands Twin Register (NTR) that has collected information from parents and teachers of twins and triplets, and from the young twins and triplets themselves. The children were registered by their parents as newborns (Beijsterveldt et al., 2013).

Child and adolescent behavioral and emotional problems

The most common behavior problems in children and adolescents can be roughly divided into 4 different types: anxious and depressive, inattentive and hyperactivity, autism related and disruptive behavior symptoms. These symptoms can be categorized into disorders according to, for example, the Diagnostic and Statistical Manual of DSM-IV disorders (APA, 2000). However, for research focusing on the etiology of problem behavior, it can be more useful to investigate continuous measures of problem behavior. Continuous measures contain more information regarding variation than a dichotomous trait and recent studies have suggested that the latent variables underlying the disorders are dimensional (Markon, Chmielewski and Miller, 2011; Haslam, Holland and Kuppens, 2012). In this thesis, we focus on *internalizing* and *externalizing* problem behavior, mostly measured with the age appropriate versions of the Achenbach System of Empirically Based Assessment (ASEBA, 2013). The Child Behavior Check List (CBCL) assesses anxious depression, withdrawn behavior, somatic symptoms, attention problems, thought problems, social problems, aggressive behavior and rule breaking behavior. The first three clusters of symptoms comprise the INT broadband scale and the latter two the EXT broadband scale (the middle three types of problems are thus not included in the two broadband scales).

Twin studies

Twin studies comprise the most often used design to estimate the heritability of human traits. By now, there hardly is a phenotype of which the heritability has not been assessed in the classical twin design (Martin, Boomsma and Machin, 1997; van Dongen, Draisma, Martin and Boomsma, 2012). Initially, a problem with many twin studies was that they tended to be far too small; experimenters who were used to the sample sizes that are required to properly compare means, assumed that these were sufficient for the comparison of correlations between monozygotic (MZ) and dizygotic (DZ) twin pairs. In a series of papers by Jinks and Fulker (1970), and by Eaves and coworkers (Eaves, Last, Martin and Jinks, 1977; Eaves, Last, Young and Martin, 1978), power studies indicated that to obtain a reasonably accurate estimate of the degree of genetic influence on a quantitative trait, at least 200 pairs need to be studied for a trait of high heritability, whereas ten or twenty times this number is needed to study traits of intermediate or low heritability, especially if other influences such as shared family environment or genetic non-additivity (genetic dominance, epitasis) are also important (Martin, Eaves, Kearsey and Davies, 1978; Posthuma and Boomsma, 2000). For dichotomous traits, such as disease status (e.g. affected versus unaffected) in which information from a continuous measure is unavailable, even larger samples of unselected twins need to be studied - the rarer the trait the higher the number of twin pairs that needs to be included in the study (Neale, Eaves and Kendler, 1994).

These power calculations reinforced the value of the large population-based twin registers which for example existed in Scandinavia, and led to the establishment of large volunteer twin registers in other countries, including the NTR, that longitudinally follow large to very large numbers of twins and their families (Beijsterveldt et al., 2013; Willemsen et al., 2013).

Emotional and behavioral problems during childhood and adolescence are among the traits investigated in these large twin studies. The results have clearly indicated that all kinds of child and adolescent problem behaviors are partially heritable, but that estimates vary depending on the phenotype. Heritability estimates are fairly high for attention problems and EXT (between 40 and 70%) and moderate, but still substantial for INT such as anxiety and depression (in general around 30 to 60%). The common environment shared by twins explains some variation in INT (10 to 40%) in children and also some variation in attention problems and EXT (20 to 40%) in children, though during adolescence the effect of common environment mostly disappears for these phenotypes.

However, it has also been acknowledged that these estimates are not static, but are population- and time-specific "snapshots" (Kendler, 2001). It has been shown, for example, that heritability estimates of child and adolescent problem behavior can depend on factors like sex, age and the environment of the child, in other words gene x age, gene x sex and gene x environment (GxE) interaction (e.g. Bartels et al., 2004; Bergen, Gardner and Kendler, 2007; Hudziak et al., 2003; Hicks, South, DiRago, Iacono and Mcgue, 2009; Hicks, DiRago, Iacono and Mcgue, 2009). Insight into differences in the etiology of problem behavior in different circumstances is important as it can facilitate the search for specific risk factors influencing a trait as well as tailored treatment for problem behavior in childhood and adolescence.

Factors interacting with heritability can be investigated by using advanced twin models. Especially GxE has received a renewed interest after the publication by Caspi et al. (2003) showing that the most investigated genetic variant for internalizing traits, the serotonin transporter length polymorphism (5-HTTLPR), only makes a difference in depressive symptoms when individuals were exposed to negative life-events. This has led to several studies trying to replicate this effect. So far, no conclusive results can be drawn from these studies given that meta-analyses come to different conclusions, which are followed by a still ongoing debate around the differences in conclusions (Munafo, Durrant, Lewis and Flint, 2009; Caspi, Hariri, Holmes, Uher and Moffitt, 2010; Duncan and Keller, 2011; Risch et al., 2009).

One of the problems with candidate GxE interaction studies is the selection of candidate genes. The knowledge about the biological mechanisms leading to psychiatric disorders or symptoms is limited. Furthermore, the field is plagued by publication bias for positive results (Duncan and Keller, 2011). As a consequence, candidate GxE studies might be followed up on insufficient grounds, just because the negative studies were not published.

Because of these problems in the candidate GxE studies, it is, at this moment, possibly better to (re)turn to twin studies when examining GxE. Twin studies can provide knowledge on GxE without the measurement of genetic variants by analyzing whether the heritability differs in certain environments. The first GxE studies compared the heritability in a group of twins exposed to an environmental factor with the heritability in a group of twins non-exposed to the environmental factor (e.g. Boomsma, de Geus, van Baal and Koopmans, 1999). After these first studies, the twin model has been extended to also include continuous moderators on the influences of additive genetic, common environmental and unique environmental factors enabling more extensive GxE analyses (Purcell, 2002).

An important issue to consider when investigating GxE is gene-environment correlation (rge), which means that an individual's genetic make-up influences the chance of being exposed to a certain environment by, for example, creating one's own environment (Kendler, 2001; Kendler and Baker, 2007; Scarr and McCartney, 1983). When environmental variables are in fact *correlated* with the genetic effects on the trait rather than *moderating* the genetic effects on the trait (i.e. GxE), this correlation can appear as interaction in typical analyses of GxE. Including the environmental factor as a fixed effect in the means model will effectively remove any genetic effects that are shared between the trait and the moderator from the covariance model (Purcell, 2002).

There are two other points that should be noted in the light of GxE research. First, there are different mechanisms leading to GxE (Dick (2011) for a review). The best known is the diathesis-stress model, which proposes that the environment works as a *contextual trigger*, such that stressors interact with personal predispositions to produce disease states, illness, and decrements in well-being (Rende and Plomin, 1992). The bio-ecological model states the opposite, i.e. genetic expression is higher in an enriched environment and thus lower in a disadvantaged environment (Shanahan and Hofer, 2005; Bronfenbrenner and Ceci, 1994). An example of this model is the social push hypothesis, which postulates that antisocial behavior in a disadvantageous environment is more environmentally driven while in an advantageous environment it is more genetically driven (Raine, 2002). Belsky and Pluess (2009) have proposed that some individuals have a genetic make-up that is relatively indifferent to the environment leading to rather stable trait scores over different environments. Other individuals, in contrast, are relatively susceptible to protective as well as risk environmental factors. They score lower on a trait when they are exposed to the protective environment but higher when they are exposed to the risk factors, compared to the indifferent individuals.

Second, differences in heritability in different environments cannot only be explained by GxE, but also by common environment x unique environment or unique environment by unique environment interaction. This is because the heritability is a ratio of the genetic variance and the total variance and thus can differ due to differences in the genetic variance as well as in the total variance. If the total variance is, for example, increased due to an increase in unique environmental variance, the heritability is decreased, even if the genetic variance remained the same.

Twin and higher-order multiples

The statistical power studies mentioned above also indicated the value of collecting data in relatives of twins (e.g. their parents and siblings) and also hint at the value of including higher-order multiples such as triplets and quadruplets. The birth of high-order multiples used to be a rare (or very rare) event, but this changed with the introduction of artificial fertility enhancing techniques. Before the introduction of In Vitro Fertilization (IVF) the rate at which triplets were born was the square of the twinning frequency in a population (e.g. if the twinning frequency is 1 in 80 births, the triplet frequency is 1 in 6400 births; Hellin, 1895; Fellman and Eriksson, 2009b; Fellman and Eriksson, 2009a). Twins can be MZ or DZ (born from one or two fertilized eggs), multiple-birth siblings are either MZ or polyzygotic; for example triplets can be mono-, di- or trizygotic. Most triplets born after Assisted Reproductive Technology (ART) are trizygotic.

Outline of this thesis

The NTR has recruited twins and triplets since its start in 1987 and now includes over 30,000 families with twins enrolled at birth and 677 families with triplets. In January 2013, 1756 persons from triplet families are active participants. A recent overview of the data collection procedures and the recruitment of families is given in van Beijsterveldt et al. (2013) which summarizes 25 years of NTR research in young twins and triplets.

The first three studies in this thesis use twin data on problem behavior collected between age 3 and age 16 and focus on gene x age, gene x sex and GxE interaction. Chapter 2 investigates differences in the impact of additive genetic influences (A), common environment (C) and unique environment (E) at ages 12, 14 and 16 years in a bivariate model of anxious depression and withdrawn behavior. Chapters 3 and 4 both examine GxE interaction. In chapter 3, teacher ratings of INT and EXT behavior of 7-, 10 and 12-year-old twins were analyzed. Heritability estimates were compared between groups of twins that were rated by the same teacher or by different teachers. If there would be differences in these heritability estimates, these could constitute evidence for gene x teacher/classroom interaction. In chapter 4, the effect of socio-economic status (SES) and formal child care attendance on the influences of A, C and E are investigated for INT and EXT at the ages of 3, 5 and 7 years.

The last two studies turn to the analysis of data from triplets. The average gestational age of triplets is described in chapter 5 of this thesis. Gestational age is an important determinant of birth weight, and triplets are often born with a low birth weight. Chapter 5 describes the association of gestational age with birth weight and also assesses the influence of other factors, including genetic factors on birth weight in triplets. Next, chapter 6 describes a unique study in triplets that additionally assessed the effects of chorionicity on triplet birth weight. This assessment was possible because the triplet data from the NTR were linked to the data on chorionicity from PALGA, the nationwide network and registry of histo- and cytopathology in the Netherlands (Casparie et al., 2007). PALGA was founded in 1971 and has nationwide coverage since 1991. In addition, to investigate the representativeness of the sample, the data from triplets registered with the NTR were compared with the total Dutch triplet population, based on information from the Netherlands Perinatal Registry (PRN-foundation).

CHAPTER 2:

HERITABILITY OF ANXIOUS-DEPRESSIVE AND WITHDRAWN BEHAVIOR: AGE-RELATED CHANGES DURING ADOLESCENCE

This chapter is published as:
Lamb DJ, Middeldorp CM, van Beijsterveldt CEM, Bartels M, van der Aa N, Polderman TJC, Boomsma DI. (2010). Heritability of anxious-depressive and withdrawn behavior: age-related changes during adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry; 49(3), 248 – 255.*

Abstract

Objective: To explain the differential course of anxiety and depression in individuals from childhood to adulthood by examining age-related changes in the genetic and environmental etiology of anxious and depressive symptoms.

Method: A sample of 1470, 1839 and 2023 Dutch twins aged 12, 14 and 16 years reported on symptoms of anxious depression (AD) and withdrawn behavior (WB), using the Youth Self Report (YSR). AD and WB were analyzed with bivariate cross-sectional genetic models for each age group to obtain estimates of the relative influence of genes (A), shared (C) and non-shared (E) environment.

Results: The best-fitting models revealed no difference between heritability estimates in boys and girls. Familial clustering at age 12 was explained by genetic and shared environmental factors. At ages 14 and 16 years, genetic factors were sufficient to explain familial clustering, shared environmental effects were absent. Genetic influences on AD and WB correlated highly.

Conclusions: These findings are in agreement with earlier studies on age-specific effects of genes and shared environment on anxiety, depression and withdrawn behavior in childhood and adolescence. The current study demonstrated that the decrease in the role of shared environment occurs after age 12. Hormonal changes accompanying the onset of puberty do not seem to explain the change in risk factors, as in 90% of the subjects, puberty had already started. More knowledge on age-specific risk factors may offer opportunities for therapeutic interventions.

Introduction

Anxiety disorders and depression are common disorders throughout the lifespan. Prevalence studies have established that between ages 5 and 17 around 8% of children suffers from an anxiety disorder. Nearly 5% in this age range suffers from major depression, but the prevalence tends to increase with age (peaking in adolescence)(Costello, Egger and Angold, 2005). In adult population-based studies, lifetime prevalence estimates for any anxiety disorder (including panic disorder, phobias, obsessive compulsive disorder, and generalized anxiety disorder) and major depression are around 19% and 16%, respectively (Bijl, Ravelli and van Zessen, 1998; Kessler et al., 2005).

Longitudinal studies have established that anxiety disorders during childhood are moderately stable, and are predictive of other mental health problems, especially depression, in later life (see Rapee, Schniering and Hudson (2009) for a review). However, these studies have also shown that childhood anxiety disorders may disappear over time without further sequelae, and that anxiety disorders may present themselves in later life, i.e., after adolescence (Rapee et al., 2009). Similar results hold for major depression, although the effects of childhood or adolescent depression seems to be more deleterious (Dunn and Goodyer, 2006; Pine, Cohen, Gurley, Brook and Ma, 1998). Co-morbidity within anxiety disorders and co-morbidity of anxiety disorders with depression is frequent in both children and adults (Angold, Costello and Erkanli, 1999; Brady and Kendall, 1992; Costello, Mustillo, Erkanli, Keeler and Angold, 2003; Ford, Goodman and Meltzer, 2003; Kessler, 1995). Research suggests that this may be due to common, mostly genetic, risk factors (Middeldorp, Cath, van Dyck and Boomsma, 2005).

Changes in the influence of genetic and environmental risk factors with age explain in part the differential course of anxiety and depression in individuals from childhood to adulthood. Specifically, reviews have established that genetic factors contribute to individual differences in anxiety and depression in both childhood and adulthood. However, the contribution of shared environmental factors is limited to childhood (Rapee et al., 2009; Gregory and Eley, 2007; Hettema, Neale and Kendler, 2001; Rice, Harold and Thapar, 2002a; Rice, 2009; Sullivan, Neale and Kendler, 2000; Middeldorp and Boomsma, 2009).

Studies performed on anxiety and depression in children and adolescents, however, have not yielded consistent results. Most studies in which the samples were divided into two or more age groups have tended to find that familial resemblance is due to genetic and shared environmental factors in childhood, whereas in adolescence genetic factors are important but shared environmental factors are not (Thapar and McGuffin, 1994; Feigon, Waldman, Levy and Hay, 2001; Kendler et al., 2008; Rice, Harold and Thapar, 2002b; Schmitz, Fulker and Mrazek, 1995; Scourfield et al., 2003). Associated with the decrease in shared environmental effects is a general increase in heritability of anxiety and depression. This finding is supported by a meta-analysis (Bergen et al., 2007). However, the results are inconsistent in that studies have also reported a decreased role of genetic and an increased role of shared environmental factors in the development from childhood to adolescence (Gjone, Stevenson, Sundet and Eilertsen, 1996). In addition, some studies have reported distinct changes in boys and girls, whereas other studies reported no clear changes at all (Legrand, Mcgue and Iacono, 1999; Eley and Stevenson, 1999; Silberg et al., 1999; Kendler, Gardner and Lichtenstein, 2008).

This inconsistency may be due to the use of different instruments to measure the phenotypes of interest. In addition, the measures differ with respect to informant or rater (i.e., the parent, the teacher, or the child itself). Reviews of studies of anxiety and depression have suggested that estimates of heritability based on parental ratings are higher than the heritabilities based on child self-report (Rapee et al., 2009; Gregory and Eley, 2007; Rice et al., 2002a). In addition, estimates of shared environmental influences tend to be lower when based on parental ratings.

Another source of inconsistency may be the variation in the age range of the participants. Almost all of the studies divided the children into two age cohorts: below and above 11 or 12 years of age. Four studies were exceptions, of which three found results that differed from those of most other studies (Kendler et al., 2008; Gjone et al., 1996; Kendler et al., 2008; Legrand et al., 1999). Studies performed in children at ages of 3, 7, 10 and 12 indicated that the effect of the common environment is still present at age 12 for different measures of anxiety and depression (Bartels et al., 2004; Hoekstra, Bartels, Hudziak, van Beijsterveldt and Boomsma, 2008; Boomsma, van Beijsterveldt and Hudziak, 2005). Thus the standard cut-off age of 11 or 12 may produce an age group that is actually heterogeneous in the sense that common environmental effects may be present in the 12 year olds, but not (or to a lesser degree) in children over 12. Thus, it is important to retain an appropriate age resolution in the study of age changes in the effects of genetic and environmental influences on anxiety and depression.

The present aim is to study the course of genetic and environmental influences on anxiety and depression at 12, 14, and 16 years. In a population based sample of monozygotic (MZ) and dizygotic (DZ) Dutch adolescent twins, anxious depression (AD) and withdrawn behavior (WB) were measured by self-report at these ages. AD and WB are subscales of the Youth Self Report (YSR; Achenbach and Rescorla, 2001), and are highly related to DSM-IV diagnoses of anxiety disorders and depression (APA, 2000; Doyle, Mick and Biederman, 2007). At each age, a bivariate genetic model was used to estimate the common and phenotype specific influences of genetic (A), shared environmental (C) and unique, non-shared, environmental (E) factors on AD and WB. To address the issue of possible influence of raters on the outcomes, we compare results with those of the analyses of the parental reports on AD and WB for this sample at age 12 (Hoekstra et al., 2008; Boomsma et al., 2005).

Methods

Subjects

All twins were registered at birth by their parents with the Netherlands Twin Register (NTR). Parents and teachers completed questionnaires concerning the twins' behavior up to age 12. In this study, the participating twins were born between 1987 and 1996 and completed the YSR at age 12, 14 and 16. Starting in 2004, after having ascertained parental informed consent, questionnaires were sent to the adolescent twins at ages 14 and 16 years. In addition to the ongoing data collection at the NTR, sub-samples of twins are also invited to participate in other projects. Within these projects, YSR data were obtained of twins at age 12. Of these twins, 340 participated in a project on behavior problems, cognition, and hormones. 806 twins participated in a project on attention problems and hyperactivity, and 392 twins participated in a project on attention and executive functioning. (Bartels, 2002; Derks, 2006; Polderman et al., 2006). The 12-year olds that took part in the project on attention problems and hyperactivity, were selected on the basis of either low or high scores on attention problems. Selection of the other 12-year-olds was based only on place of residence (at reasonable travel distance from Amsterdam). For a more detailed description of the data collection and the determination of zygosity, see Bartels et al. (2007) and Boomsma et al. (2006).

Some overlap exists between the groups. Of all twins who reported on the YSR at age 12, 39% returned the list at age 14, and 21% returned the YSR at age 16. There is no overlap in the samples of twins, who returned the YSR at age 14 and 16.

Instrument

AD and WB data came from the YSR (a self-report questionnaire). The YSR is one of the instruments in the Achenbach System of Empirically Based Assessment (ASEBA). Each ASEBA instrument contains 113 behavioral items with a 3-point scale response format. The AD subscale comprises 16 items, with possible responses 0 (= not at all), 1 (= a little bit or sometimes), and 2 (= a lot or most of the time). The range of the score on the AD subscale of the YSR is therefore from 0 to 32. Two examples of items on this subscale are "I feel lonely" and "I cry a lot". The WB subscale comprises 7 items, and given the identical response format and scoring, its range is from 0 to 14. Two examples of items of this subscale are "I am rather alone than with others" and "I refuse to talk".

Statistical Analyses

Because the data of both subscales were positively skewed (L-shaped), we used a threshold model to analyze individual differences in the liability to AD and WB. Use of the threshold model requires the discretization of the observed distributions, i.e., the creation of scores that are amenable to the analysis (it is computationally impractical to retain the original scales of 0 to 14 and 0 to 32). Therefore at each age group, AD and WB scores were categorized into three groups with roughly an equal number of subjects. In this manner, we arrived at ordinal phenotypic scores that were coded 0, 1, and 2. In the liability-threshold model, we require two thresholds to model the three categories of AD, and two thresholds to model the three categories of WB. We estimated the thresholds separately in boys and girls to take into account sex differences in AD and WB scores.

The basic assumption underlying the threshold model is that a latent continuous liability underlies the observed categorical variables. The liability is assumed to be standard normally distributed (i.e., mean of zero, standard deviation of one). Individual differences in the liability are modeled to be due to genetic and environmental effects. The adoption of the threshold model is a generally accepted way to analyze L-shaped distributed data (Derks, Dolan and Boomsma, 2004).

In each age group, we analyzed the MZ and DZ twin data simultaneously in a bivariate model (i.e., including AD and WB). We first fitted a saturated model, in which we estimated polychoric correlations within twin pairs within and across traits allowing differences between sexes and zygosity groups. We used the software package Mx to this end (Neale M.C., Boker S.M. and Maes, 2006).

Estimating correlations for MZ and DZ twin pairs constituted a first step towards evaluating the relative influence of genetic and environmental factors on trait variances and covariances between traits. MZ twin pairs are genetically identical, whereas DZ twin pairs and non-twin siblings share, on average, 50% of their genetic material. If the MZ twin correlation is higher than the DZ correlation for a certain trait (i.e. within-trait correlation) or between two traits (i.e. cross-twin cross-trait correlation), we infer that genetic factors (A) contribute to the phenotypic variances and covariance of two phenotypic traits. If DZ correlations are higher than half the MZ correlation, we infer that environmental effects (C) shared by members of the same family are present. Finally, we attribute to non-shared environment effects (E) variance that is not due to genetic or shared environmental factors. These are environmental influences that are not shared by family members. The non-shared environmental variance component also includes measurement error variance.

Next, using structural equation modeling in Mx, we fitted genetic bivariate models to the data. A graphical representation of the genetic model is given in Figure 1. The amount of variance in a single trait and covariance between traits due to A, C and E can be calculated from the factor loadings a, c and e. Parameter estimates for a, c, and e were allowed to differ with respect to sex. The results of this model allow us to breakdown the phenotypic correlation between AD and WD into genetic and environmental correlations.

We performed the analyses using the raw-data likelihood estimation in Mx. We fitted various models which were nested in the sense that one model could be derived from the other by the imposition of one or more constraints on the parameters. We evaluated the differences between (nested) models by inspecting the log-likelihood ratio test (LRT) statistic. This is minus twice the difference of the maximum log-likelihoods of two models. Asymptotically this statistic has a χ^2 distribution, with the degrees of freedom (df) equaling the difference in df between the two models. If the p-value associated with this statistic is greater than our chosen alpha of 0.05 we concluded that the constraints associated with the more parsimonious model are tenable.

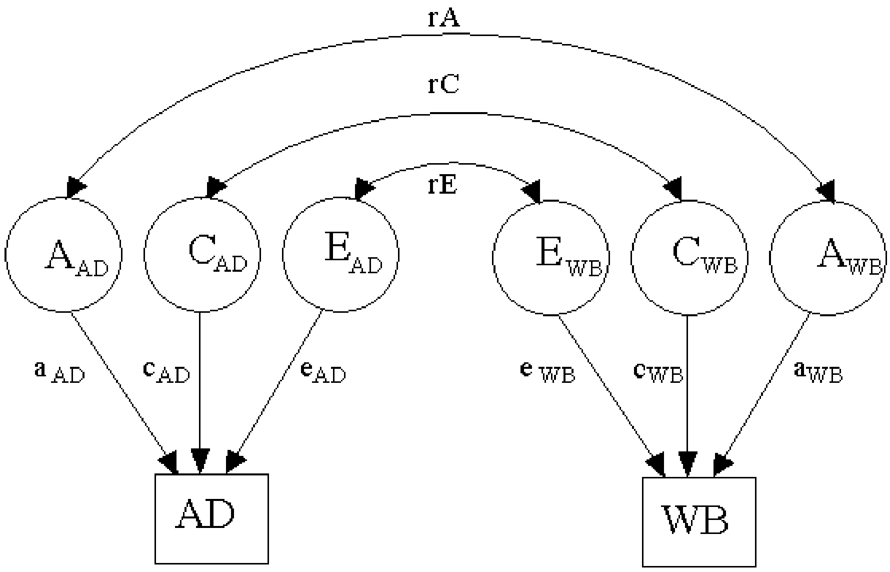


Figure 1: Path diagram of the bivariate genetic model. The observed variables are represented with a square (AD = anxious depressive, WB = withdrawn behavior), the latent factors by circles (A = additive genetic, C = shared environment, E = non-shared environment). The variance within an observed variable is broken down into the three latent factors, and the influence of the latent factors is given by the path coefficients a, c and e. The total variance of an observed variable is sum of the squared path coefficients (for AD: $a_{AD}^2 + c_{AD}^2 + e_{AD}^2$). The covariance between the two observed variables is also broken down into a genetic, shared, and nonshared environmental part. The phenotypic correlation is the standardized covariance between the two observed variables; this can be expressed as $rP = a_{AD}\ rA\ a_{WB} + c_{AD}\ rC\ c_{WB} + e_{AD}\ rE\ e_{WB}$. *Note:* A = additive genetic; C = shared environmental; E = nonshared environmental

First, sex differences in the relative influence of A, C, and E were tested by constraining the parameter estimates a, c, and e over sex to be equal, and testing whether this constraint led to decrease in goodness of fit. The total variance of the liabilities underlying WB and AD were constrained at 1 by specifying that var(E) = 1 – (var(A) + var(C)). Consequently, when testing for sex differences in the genetic architecture the test has 7 df. Second, the statistical significance of the genetic and environmental parameters was tested by constraining them at 0.

Results

Descriptives

The number of complete and incomplete twin pairs as a function of age and zygosity is given in Table 1. At ages 12, 14 and 16 years, the total sample consisted of 726 boys and 744 girls, 801 boys and 1038 girls, and 903 boys and 1120 girls, respectively. The mean age in years in each age group was 11.9 (SD = 0.52, range 11 to 13), 14.1 (SD = 0.41), and 16.7 (SD = 0.90, range 16 to 19). Figure 2 shows the mean scores on AD and WB as a function of age and sex. Girls scored higher than boys and, in general, scores tend to increase with age. In view of these results, thresholds were estimated separately per age group and sex. Phenotypic correlations between AD and WB were high in both sexes at all ages, ranging from 0.59 to 0.69.

Genetic analyses

Twin correlations as a function of age, sex and zygosity are shown in Table 1. At each age, MZ correlations were higher than DZ correlations, suggesting that genetic factors contribute to individual differences in AD and WB. However, the MZ correlations at age 12 were less than twice as high as the DZ correlations, indicating that at that age shared environmental influences may be present. The cross-trait-cross-twin correlations showed largely the same patterns.

Fit statistics of the three full genetic models and the nested submodels can be found in Table 2. There were no significant sex differences in the effect of A, C and E at any age (all p values > .05). At age 12, neither effect of A nor C could be dropped from the model ($\chi2(3)$ = 10.4, p = .02, and $\chi2(3)$ = 20.6, p < .01, respectively). This means that at age 12 both genetic and shared environmental effects contribute to the phenotypic variance in AD and WB. In AD, 35% of the variance was explained by A, and 21% by C, in WB these estimates were 3% and 38%, respectively. At age 14 and 16, the C component could be dropped from the model ($\chi2(3)$ = 0, p = 1.00, and $\chi2(3)$ = 0.8, p = .84). The influence of the genetic factors ranged from 37% to 67% on AD and WB at age 14 and 16. Table 3 contains the estimates of the relative influence of genes, shared, and non-shared environment, with the associated confidence intervals (CI).

The YSR AD and WB scales are highly correlated. This correlation is mostly explained by shared genes. The correlations between genetic factors are 1.0 at age 12, and 0.85 at age 14 and 16. At age 12 the correlation between shared environment influencing AD and WB was also 1.0. Unique environmental factors were less correlated, with estimates around 0.5.

Table 1. *Twin correlation for the categorized anxious depressed and the withdrawn behavior scores for age 12, 14 and 16, and the cross-twin cross-trait correlations of AD and WB.*

	AGE 12					AGE 14					AGE 16				
Zygosity	CP (n)	IP (n)	AD	Cross	WB	CP (n)	IP (n)	AD	Cross	WB	CP (n)	IP (n)	AD	Cross	WB
MZM	140	2	0.57	0.41	0.38	130	3	0.53	0.30	0.37	175	13	0.49	0.35	0.44
DZM	138	3	0.39	0.39	0.38	125	7	0.28	0.23	0.23	132	16	0.28	0.25	0.10
MZF	161	3	0.58	0.34	0.46	222	7	0.76	0.51	0.42	215	14	0.57	0.45	0.53
DZF	124	6	0.47	0.45	0.38	143	11	0.42	0.15	0.03	187	14	0.23	0.19	0.18
DOS	162	6	0.17	0.23	0.36	271	29	0.29	0.22	0.17	241	66	0.40	0.28	0.23

Note: AD = anxious depressed; CP = complete pairs; DOS = dizygotic twins of opposite sex; DZF = dizygotic female; DZM = dizygotic male; IP = incomplete pairs; MZF = monozygotic female; MZM = monozygotic male; WB = withdrawn behavior.

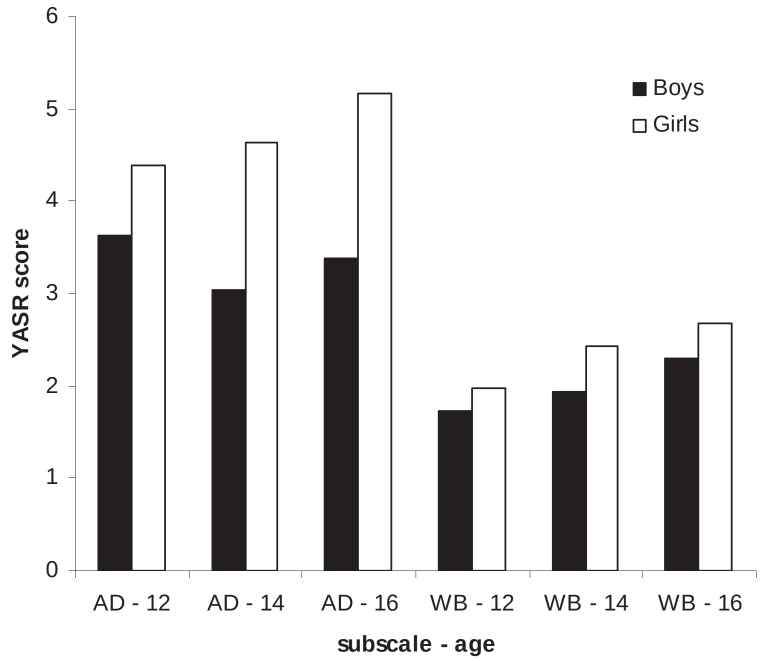


Figure 2: Mean scores on anxious depressed and withdrawn behavior as a function of sex and a

anxious depressed; WB = withdrawn behavior; YASR = Young Adult Self Report.

Table 2. *Model fitting results for the bivariate genetic analyses on anxious depression and withdrawn behavior scores.*

	Model	-2LL	Compared with model	χ^2	Δdf	*p*	AIC
Age 12	0. SAT	5785.7					-74.3
	1. ACE	5794.1	0	8.4	3	.04	-71.9
	2. ACE sex equal	**5804.7**	**1**	**10.7**	**7**	**.15**	**-79.3**
	3. AE	5815.2	2	10.4	3	.02	-74.8
	4. CE	5825.3	2	20.6	3	.00	-64.7
Age 14	0. SAT	7313.6					-46.4
	1. ACE	7317.4	0	4.1	3	.25	-48.6
	2. ACE sex equal	7331.3	1	13.9	7	.05	-52.7
	3. AE	**7331.3**	**2**	**0**	**3**	**1.00**	**-58.7**
Age 16	0. SAT	7903.4					-210.5
	1. ACE	7911.3	0	7.8	3	.05	-208.7
	2. ACE sex equal	7917.1	1	5.7	7	.57	-220.9
	3. AE	**7917.9**	**2**	**0.8**	**3**	**.84**	**-226.1**

Note: The best model fit is displayed in boldface type. A = additive genetic; AIC = Akaike's Information Criterion; C = common environment; E = unique environment; SAT = saturated model.

Table 3. *Standardized estimates of amount of variance explained by Additive genetic (A), Common (C), and unique Environmental (E) influences and estimates of correlations among these influences from bivariate genetic analyses.*

Age	A_{AD}	C_{AD}	E_{AD}	A_{WB}	C_{WB}	E_{WB}	r_A	r_C	r_E
12	0.35	0.21	0.45	0.03	0.38	0.60	1.00	1.00	0.50
	(0.25 – 0.44)	(0.09 – 0.30)	(0.37 – 0.56)	(0.00 – 0.10)	(0.25 – 0.47)	(0.51 – 0.68)	(0.00 – 1.00)	(1.00 – 1.00)	(0.24 – 0.59)
14	0.67	-	0.33	0.37	-	0.63	0.85	-	0.42
	(0.46 – 0.72)		(0.26 – 0.41)	(0.18 – 0.47)		(0.53 – 0.68)	(0.75 – 0.97)		(0.30 – 0.48)
16	0.55	-	0.45	0.45	-	0.55	0.85	-	0.51
	(0.46 – 0.63)		(0.37 – 0.54)	(0.35 – 0.54)		(0.46 – 0.65)	(0.75 – 0.94)		(0.40 – 0.60)

Note: AD = anxious depressed; WB = withdrawn behavior.

Discussion

The aim of present study was to provide insight into the age specific genetic and environmental influences on anxiety and depression during adolescence. Our results indicate that at age 12 familial clustering of AD and WB was due to genetic and shared environmental risk factors, whereas at age 14 and 16, familial clustering was due only to genetic factors. We found no change in the extent to which unique environmental factors play a role from age 12 to age 16, with estimates of explained variance around 0.4 for AD and around 0.55 for WB. The age-related changes in the effects of A and C are in agreement with most other studies that addressed differences between childhood and adolescence in the genetic and environmental effects on anxiety and depression (Feigon et al., 2001; Kendler et al., 2008; Rice et al., 2002b; Schmitz et al., 1995; Scourfield et al., 2003; Thapar and McGuffin, 1994).

Because of the narrow age range of the groups, we could establish that the decrease in the effect of C occurs after age 12. A similar decrease in the influence of C has also been reported with respect to obsessive compulsive symptoms (van Grootheest et al., 2008). Several explanations are possible for the change in the relative contributions of risk factors for anxiety and depression in this period in life. One is that the onset of puberty and hormonal changes lead to changes in gene expression influencing the risk for anxiety and depression, i.e. new genes come to expression or other genes come to expression at a higher level. In this case, the change in risk factors of AD and WB would be more or less in synchrony with the start of puberty. We investigated this hypothesis by using self-report data inquiring about pubertal physical changes. Genetic analyses indicated that the decrease in C occurred after age 12. Therefore, one would expect that, if the hormonal changes responsible for the start of puberty play a role in the decrease in the influence of C and the increase in the influence of A, the majority of the 12-year-old twins would be prepubescent. However, this was not the case: based on self-report, only around 10% of the 12-year-old twins appeared to be in Tanner stage 1 (no pubic hair or breast development) (Marshall and Tanner, 1969; Marshall and Tanner, 1970). Thus, in 90% of the present 12-year-olds, puberty had already started. Another explanation is that the influence of C decreases due to a gradual attenuation of the level of parental control and time spent at home and together with the co-twin. This hypothesis requires further investigation.

As the results of similar analyses relating to parental ratings of offspring at 12 on both AD and WB are available, we can compare these results with our present results. Our present estimates of the effects of A and C on AD and WB at age 12 are similar to those obtained using parental reports (Bartels et al., 2004; Hoekstra et al., 2008; Boomsma et al., 2005). Only with respect to WB was the present estimate of the effect of C higher than reported by Hoekstra et al., who found very little influence of C in the maternal ratings and an influence on the parental ratings amounting to approximately 15% of the variance. These results are consistent with the results of earlier studies that used parental ratings. In these studies the effect of C were small or absent (Rapee et al., 2009; Gregory and Eley, 2007; Rice et al., 2002a).

Discrepancies with results of other studies can be due to variation in the age range of the participants in the studies. In most other studies the age ranges are appreciably broader than the present age range (Eley and Stevenson, 1999; Silberg et al., 1999). However, this does not explain why Kendler et al. (2008) and Legrand et al. (1999) found no age differences in the es-

timates of A and C, whereas Gjone et al. (1996) found an increase in the effect of C with age. Kendler et al. did not actually test the effects of C by age, so they may have missed some effect. They also used parental reporting in addition to self-reporting. There is no clear explanation for the lack of age specific effects found by Legrand et al.. There is also no clear explanation for the effect of C at age 14 to 15 found by Gjone et al. in a Norwegian twin sample. However, at least compared to the present Dutch samples, cultural differences may play a role.
We found no sex effects on heritability estimates of AD and WB. The literature is mixed on this point, with most studies reporting higher heritability estimates in girls, although higher heritability estimates in boys and the general absence of sex effects are also reported (Feigon et al., 2001; Scourfield et al., 2003; Kendler et al., 2008; Eley and Stevenson, 1999; Silberg et al., 1999). This discrepancy might be due to differences in power and heterogeneity in phenotypes.

A limitation of the present study is that the overlap in samples was too small to perform longitudinal analyses. In a longitudinal analysis, it is also possible to investigate whether similar or different genetic or environmental factors influence AD and WB over time. Analyses from the longitudinal Swedish Twin Study showed that the genetic factors influencing anxiety and depression differ between childhood and adolescence. The effect of the genetic factors influencing anxiety and depression during childhood is diminished and new genetic factors play a more important role (Kendler et al., 2008; Kendler et al., 2008; Kendler, Gardner, Annas and Lichtenstein, 2008). In the future, it would be of interest to carry out longitudinal analyses to investigate the genetic and environmental correlation between phenotypes within an age group and over time.

Overall, our study showed that influences of A and C differ over time. This difference in risk factors from childhood to adolescence can partly explain why some individuals suffer from anxiety and/or depression early in life, while others do not have any complaints during childhood, but begin to experience symptoms during adolescence. Further research is needed to identify the age-specific risk factors, either genetic or environmental, for AD and WB. More knowledge on the age-specific risk factors might facilitate the development of different therapeutic interventions for children of different ages.

CHAPTER 3:

GENE-ENVIRONMENT INTERACTION IN TEACHER-RATED INTERNALIZING AND EXTERNALIZING PROBLEM BEHAVIOR IN 7- TO 12-YEAR-OLD TWINS

This chapter is published as:
Lamb DJ, Middeldorp CM, van Beijsterveldt CE, Boomsma DI. (2012). Gene-environment interaction in teacher-rated internalizing and externalizing problem behavior in 7- to 12-year-old twins. *Journal of Child Psychology and Psychiatry; 53(8), 818 – 825.*

Abstract

Background: Internalizing and externalizing problem behavior at school can have major consequences for a child and is predictive for disorders later in life. Teacher ratings are important to assess internalizing and externalizing problems at school. Genetic epidemiological studies on teacher rated problem behavior are relatively scarce and the reported heritability estimates differ widely. A unique feature of teacher ratings of twins is that some pairs are rated by different and others are rated by the same teacher. This offers the opportunity to assess gene-environment interaction.

Methods: Teacher ratings of 3502 7-year-old, 3134 10-year-old and 2193 12-year-old twin pairs were analyzed with structural equation modelling. About 60% of the twin pairs were rated by the same teacher. Twin correlations and the heritability of internalizing and externalizing behavior were estimated, separately for pairs rated by the same and different teachers. Socio-economic status and externalizing behavior at age 3 were included as covariates.

Results*:* Twin correlations and heritability estimates were higher when twin pairs were in the same class and rated by the same teacher than when pairs were rated by different teachers. These differences could not be explained by twin confusion or rater bias. When twins were rated by the same teacher, heritability estimates were about 70% for internalizing problems and around 80% in boys and 70% in girls for externalizing problems. When twins were rated by different teachers, heritability estimates for internalizing problems were around 30% and for externalizing problems around 50%.

Conclusions: Exposure to different teachers during childhood may affect the heritability of internalizing and externalizing behavior at school. This finding points to gene-environment interaction and is important for the understanding of childhood problem behavior. In addition, it could imply an opportunity for interventions at school.

Introduction

Internalizing (INT) and externalizing (EXT) problem behavior at school can have major consequences for a child. It can seriously affect learning achievements and lead to bullying by other children or negative interactions with the teachers(Deater-Deckard, 2001; Ladd and Burgess, 1999). Moreover, a DSM-IV diagnosis requires that problem behavior during childhood is not restricted to one environment (APA, 2000). It is therefore common practice to take into account INT and EXT behavior at school as rated by the teacher when assessing a child for psychiatric problems in a child and adolescent psychiatry clinic.

Behavior genetic studies that investigate the relative influence of genetic and environmental risk factors on teacher rated behavior are relatively scarce and sample sizes are small. An often used paradigm within behavior genetics is the twins design, which makes use of the different genetic relatedness of monozygotic (MZ) and dizygotic (DZ) twins. As MZ twins are (nearly) genetically identical whereas DZ twins share on average 50% of their genes a higher correlation in MZ twin pairs than in DZ twin pairs suggest that part of the individual differences can be explained by genetic risk factors (Boomsma, Busjahn and Peltonen, 2002). A correlation in MZ twin pairs, which is less than twice the DZ twin correlation, indicates the influence of common environmental factors shared by children in the same family. Differences in problem behavior within MZ twin pairs are due to unique environmental influences, which also include measurement error.

Heritability estimates based on teacher ratings of twins behavior vary widely, e.g. from 31% to 81% for EXT and conduct problems respectively (Saudino, Ronald and Plomin, 2005; Towers et al., 2000). One of the reasons for the varying estimates is that twin correlations tend to be higher when the twins' behavior is rated by the same teacher than when it is rated by different teachers (e.g. Saudino et al., 2005; Simonoff et al., 1998; Towers et al., 2000). Assuming that same teacher ratings are inflated or that using multiple teachers to rate a twin pair increases the error in the measurement, these ratings were sometimes excluded, resulting in a wide range of heritability estimates for problem behavior (Hartman, Rhee, Willcutt and Pennington, 2007; Towers et al., 2000). Heritability is estimated around 72% for different types of problem behavior when twin pairs are rated by the same teacher, in contrast to estimates of about 54% when twin pairs are rated by different teachers (Saudino et al., 2005).

However, it is an important question whether teacher ratings are inflated or biased and if this bias explains the differences in correlations between same and different teacher ratings. In addition to teacher bias leading to correlated errors, the higher correlations in MZ twins may be explained by confusion of twins rated by the same teacher, or by gene-environment interaction (Baker, Jacobson, Raine, Lozano and Bezdjian, 2007; Derks, Hudziak, van Beijsterveldt, Dolan and Boomsma, 2006; Derks, Dolan, Hudziak, Neale and Boomsma, 2007; Polderman, Posthuma, De Sonneville, Verhulst and Boomsma, 2006; Simonoff et al., 1998; Saudino et al., 2005). Rater bias refers to the teacher's own bias which affects the similarity of ratings of twin behavior if the same teacher rates more than one child (correlated errors). Twin confusion is present when a teacher does not always distinguish the two individuals in a twin pair and therefore rates their behavior as more similar. A third explanation is the presence of gene-environment interaction (GxE). GxE implies that the influence of genetic factors depends on the environment (Boomsma and Martin, 2002; Eaves, 1984). Thus, heritability estimates can differ between groups that are rated in different environments.

In this report we focus on the etiology of teacher-rated INT and EXT problem behavior in a large sample of male and female twins at age 7, 10 and 12 years. The sample includes twin pairs who are rated by the same teacher and pairs who are rated by different teachers. First, we investigate the influence of social-economic status (SES) and earlier problem behavior on the chance of twins being separated or not. Earlier research has shown that SES and earlier problem behavior are associated with problem behavior at later ages (Bartels et al., 2004; Bradley and Corwyn, 2002). If these factors are also related to twins being separated, this can lead to confounding. Consequently, significant risk factors are included in the following analyses as covariates. Next, we tested which of the three models (twin confusion, correlated errors or GxE) described the data best, and estimated to what extent individual differences in teacher-rated INT and EXT behavior are explained by genetic and environmental risk factors.

Methods

Subjects

All twins participated in the ongoing longitudinal study of the Netherlands Twin Register (NTR; (Boomsma et al., 2002). Most twins were registered with the NTR at birth. Parents of twins completed questionnaires assessing the twin's behavior from birth to age 12. If parents provided informed consent, teachers were asked to report on the twin's behavior at age 7, 10 and 12. Assessment took place starting from 1999 up to the year 2010. The current study included twins with maternal ratings on externalizing problems at age 3, in addition to teacher ratings at age 7, 10 or 12. Teacher ratings were available for 3700 complete and 701 incomplete (i.e.

only one teacher returned the questionnaire for one twin from a pair) 7-year-old twin pairs and 3310 complete and 776 incomplete 10-year-old twin pairs, 2315 complete and 618 incomplete 12-year-old twin pairs. Twins were excluded from the analyses when teachers reported that they did not know the student well (at age 7, 10 and 12: 0.9%, 1.5%, 0.8%), or had only known the student for a short time (at age 7, 10 and 12: 1.6%, 2.0%, 1.7%). Twin pairs were excluded when one or both twins attended special education (at age 7, 10 and 12: 0.8%, 1.7% and 1.9%). Special education included schools for physical and cognitive disabled children and children with learning and extreme behavioral problems. Classes in special education are different from regular primary school classes in that they are smaller. Zygosity determination took place in different projects in which subgroups of the twins participated. For 2% of the twins, zygosity was determined based on blood group polymorphisms, for 12% based on DNA polymorphisms, for 55% based on parental ratings of zygosity and for 31% based on the fact that the twins were of opposite sex. Twins were only included if information on zygosity and SES was available leading to missing data around 1%.

Measures

Teachers completed the Teacher Report Form (TRF; Achenbach and Rescorla, 2001; Achenbach, 1991b), which contains 120 items measuring problem behavior. The TRF includes the broad band INT scale that consists of the subscales Anxious/Depressed, Withdrawn Behavior and Somatic Complaints, and the EXT scale which consists of subscales Aggressive Behavior and Rule Breaking Behavior.

Parents reported on the twins' behavior by means of the Child Behavior Check List (CBCL; Achenbach, 1991a). The CBCL and the TRF are both part of the Achenbach System of Empirically Based Assessment (ASEBA) and are comparable. The broad band INT scale for the CBCL 2/3 consists of the Anxious/Depressed and Withdrawn Behavior subscales (and not Somatic Complaints), EXT consists of the Rule Breaking Behavior and Aggressive Behavior subscales. SES of the twin families was measured at age 3, 7 and 10 years. SES was obtained from a full description of the occupation of the parents for about two-third of the twin sample. Subsequently, SES was coded according to the standard classification of occupations (CBS, 2001). For the remaining families, SES was obtained by the EPG-classification scheme (Erikson, Goldthorpe and Portocarero, 1979), combined with information on parental education. All twins were classified into three SES levels (1 = low; 2 = middle; 3 = high).

Statistical analyses

Means model and covariates. INT and EXT problem behavior showed an L-shaped distribution. Therefore, prior to the analyses, the data were normalized in PRELIS. PRELIS makes use of a normalizing transformation that fits the inverse normal density function to the ranked data (Jöreskog and Sörbom, 2002). This did not influence the estimates of the variances.

The effect of SES and problem behavior on twin separation at age 7, 10 and 12, was analyzed performing binary logistic regression analyses in SPSS (Chicago: SPSS Inc., 2008). INT and EXT as rated by the mother and the father at age 3 and SES were used as predictors and twin pair separation at age 7, 10 and 12, was used as the dependent variable.

Structural equation modeling as implemented in the software package Mx was used to estimate means, variances and twin correlations as a function of zygosity for twin pairs rated by the same and by different teachers (Neale M.C. et al., 2006). First, a so-called saturated model was fitted to the INT and EXT data at age 7, 10 and 12 in which all these parameters were freely estimated. Means were estimated separately for twins in the same class and in different classes. Sex, SES and maternal ratings of EXT at age 3 were included as covariates on the INT and EXT scores

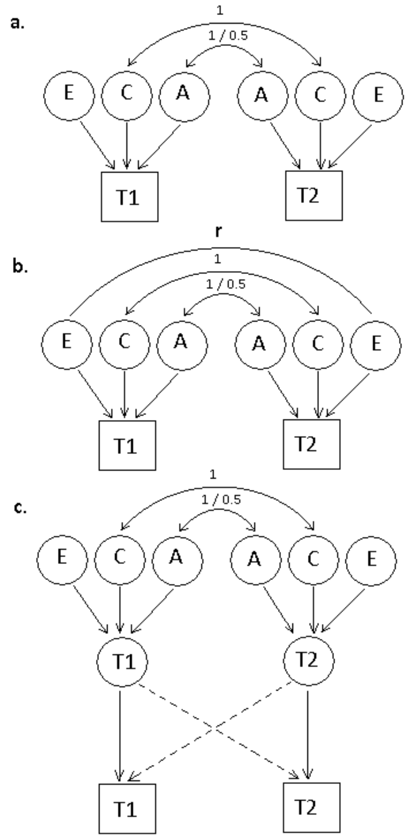


Figure 1. a. Basic genetic model to estimate the effect of additive genetic (A), common environment shared by children in the same family (C) and unique environmental (E) effects on behavioral traits. Latent variables are given in circles, observed variables in squares. The underlying assumption is that E is uncorrelated between twin 1 (T1) and twin 2 (T2), the correlation of C is 1, and A correlates in MZ twins with 1 and in DZ twins with 0.5. b. Correlated errors model; E is correlated (r) when pairs are rated by the same teacher. c. Twin confusion model; behavior of one twin is sometimes ascribed to the other twin when both are rated by the same teacher.

and were tested for significance in the saturated model. Differences in mean problem behavior between the same and different teacher groups were also tested in the saturated model.

Genetic epidemiological analyses. The classical twin design (see Figure 1a) was used to estimate the influence of genetic, shared environmental and unique environmental factors on INT and EXT data of twins rated by the same teacher and on data of twins rated by different teachers. If estimates for the genetic variance differ between same teacher ratings and different teacher ratings then this constitutes evidence for GxE. GxE can be confounded with gene-environment correlation, i.e. the phenomenon that environmental exposure is associated with individual's genetic make-up (Purcell, 2002). To correct for this effect, means were estimated separately for twins in the same class and in different classes.

Next, differences between parameter estimates for twins rated by the same and by different teachers were tested by constraining them to be equal. Further, sex differences in parameter estimates were tested as well as the significance of the effect of shared environmental factors.

Correlated errors model. Teachers can bring a bias into their ratings, which can result in children being rated as more similar when rated by one teacher. As a result, the correlations of twins rated by the same teacher are inflated. The teacher's bias can be accounted for by estimating a correlation between nonshared environment (which also includes measurement error) of twin pairs rated by the same teacher (Simonoff et al., 1998; Figure 1b).

Twin confusion model. Twin confusion is present if a teacher does not always discriminate between the two individuals in a twin pair (Simonoff et al., 1998, Figure 1c). Behavior of one twin (in the figure depicted with a circle) is sometimes ascribed to his or her co-twin (depicted with the broken line pointing to the teacher's ratings of the co-twin).

In all models sex, SES and EXT at age 3 were included as covariates on the mean.

Results

Twin pairs rated by the same teacher and by different teachers.

Table 1 shows the INT and EXT scores for the same teacher group and the different teacher group at age 3 (rated by the mother) and at age 7, 10 and 12 (rated by the teacher). The presented statistics are of the oldest twin to avoid paired observations when including both twins of a pair. EXT scores at age 3 were higher in twins who were separated at school than twins who were not separated ($p < .01$). INT scores at age 3 were only marginally higher for the separated twins ($p = .03$). A binary logistic regression showed that EXT at age 3 was positively predictive of being separated at age 7, 10 and 12 (all tests showed $p < .01$). INT at age 3 had no power to predict twin separation at later ages (all tests showed $p > .05$). Binary logistic regression also showed that SES was positively predictive of being in the different teacher group at age 7 and 10, but not at age 12 (for age 7 and 10 $p < .01$, for age 12 $p > .05$).

At age 7, same teacher ratings on INT and EXT were lower than the different teacher ratings. The same pattern was seen at age 10 and 12 (see Table 1). After correcting for the effects of sex, SES and EXT at age 3, scores on EXT at age 7 could be constrained to be equal in the same and different teacher group ($p > .05$). This was not the case for scores on EXT at age 10 and 12, and for scores on INT at all ages.

Table 1. *Raw scores on mean (SD) internalizing and externalizing behavior at age 3, 7, 10 and 12.*

		Internalizing				Externalizing			
Questionnaire		ST	DT	*t(df)*	*p*	ST	DT	*t(df)*	*p*
CBCL 2/3	Age 3	4.6 (3.8)	4.9 (4.1)	2.2 (3716)	0.03	15.6 (9.5)	17.1 (10.1)	4.8 (3490)	0.00
TRF	Age 7	4.1 (4.7)	5.1 (5.5)	6.4 (3905)	0.00	4.1 (6.8)	4.7 (7.1)	2.7 (4177)	0.01
TRF	Age 10	4.4 (5.3)	5.7 (6.1)	7.3 (3214)	0.00	4.4 (7.1)	5.3 (8.0)	3.6 (3346)	0.00
TRF	Age 12	3.8 (5.0)	5.2 (5.8)	6.3 (1791)	0.00	4.1 (7.0)	5.2 (8.1)	3.7 (1838)	0.00

Note: ST = same teacher, DT = different teacher, CBCL = Child Behavior Check List, TRF = Teacher Report Form.

The effects of sex, SES and EXT at age 3 on INT and EXT were also tested in the means model. Sex had no significant effect on INT scores at age 7, 10 or 12 (all tests showed $p > .05$). There was a significant effect of sex on EXT behavior at age 7, 10 and 12 (all tests showed $p < .01$), with boys scoring higher on EXT than girls. There was an effect of SES on mean INT and EXT scores at each age, except for INT at age 7 ($p > .05$). Mean scores on problem behavior decreased with increasing SES. EXT behavior at age 3 was a significant predictor of mean INT and EXT scores at each age (all tests showed $p < .01$).

Genetic analyses. Twin variances and correlations as a function of zygosity, age, sex and rating by the same or by different teachers are presented in Table 2. Twin pair correlations for MZ twins rated on EXT by the same teacher were higher than MZ twins rated by different teachers, the same holds for the DZ twins. Roughly the same pattern was seen for ratings on INT, with a few exceptions. MZ correlations were higher than DZ correlations, indicating that genetic factors influence problem behavior. In general, the MZ correlations were about twice as high as the DZ correlations, which indicate little or no influence of shared environmental factors. This was the case for both the same teacher and different teacher group. No influence of shared environment in the same teacher group indicates that the correlated errors model probably does not fit the data, as in this model the presence of shared environment is expected because MZ as well as DZ correlations are inflated due to correlated errors. Furthermore, the variances were about equal for the same and different teacher group, indicating that the twin confusion model is also not likely to describe the data well, as variances of the same teacher ratings are inflated when teachers confuse twins.

The correlated errors, twin confusion and the GxE interaction models were compared to the saturated model (see supplementary Table 1 for fit statistics). At all ages and for both phenotypes the GxE model gave a good fit to the data, except for INT at age 7 which showed a marginally significant deterioration of fit compared to the saturated model ($p = .03$). Both the twin confusion model and the correlated errors model did not fit the data, except for the correlated errors model on INT at age 12. Because the overall trend (a good fit of the GxE model at all ages) we adopted this model for INT and EXT.

The raw estimates and the proportions of variance explained by additive genetic, common and unique environmental factors are presented in Table 3. At all ages, heritability estimates were higher in the same teacher group than in the different teacher group. There was no effect of shared environment at any age for any type of problem behavior.

For INT there was no difference between boys and girls in the amount of variance explained by genetic or unique environmental factors. This was the case for the same teacher group and for the different teacher group. At all ages heritability estimates were about 70% in the same teacher group. Heritability in the different teacher group was nearly 30% at age 7 and 10, and 39% at age 12.

There was a significant effect of sex on the heritability estimates of EXT, although the effect was small. For boys, estimates of the effect of additive genetic factors were between 79 and 82% in the same teacher group. In the different teacher group heritability estimates were between 47 and 53%. For girls, the heritability estimates were lower in the same teacher group (67-75%) and marginally higher (48-56%) in the different teacher group.

Table 2. *Variances and correlations for internalizing (INT) and externalizing (EXT) as a function of twin pairs rated by the same or by different teachers and zygosity.*

	Age 7						Age 10						Age 12					
	Same teacher			Different teacher			Same teacher			Different teacher			Same teacher			Different teacher		
Zyg	Var	*r*	*N* pairs	Var	*r*	*N* pairs	Var	*r*	*N* pairs	Var	*r*	*N* pairs	Var	*r*	*N* pairs	Var	*r*	*N* pairs
	Internalizing																	
MZM	23	.65	315+7	25	.27	242+5	30	.72	320+4	31	.28	182+5	27	.68	242+2	28	.37	123+6
DZM	26	.33	275+5	27	.21	267+16	33	.47	273+4	32	.19	219+7	24	.39	214+7	26	.18	120+4
MZF	25	.71	380+4	26	.33	273+8	29	.66	408+4	29	.22	210+16	26	.67	349+9	29	.43	142+5
DZF	24	.20	302+9	22	.10	235 +9	30	.37	274+7	32	.22	174+17	25	.37	211+6	27	.18	106+2
DOS	24/23	.36	622+17	24/28	.08	465+18	29/28	.34	550+18	28/29	.07	410+16	26/25	.41	403+11	29/26	.20	229+9
	Externalizing																	
MZM	50	.81	324+1	52	.49	246+1	61	.78	324+2	54	.39	188+0	54	.81	244+0	54	.50	127+2
DZM	50	.37	278+3	49	.25	279+4	65	.37	279+1	69	.37	224+2	58	.44	219+2	67	.37	121+3
MZF	37	.67	388+1	39	.55	297+2	40	.75	418+3	45	.50	220+0	38	.74	355+3	40	.56	146+1
DZF	37	.34	310+3	38	.27	233+3	40	.40	280+2	44	.11	190+2	38	.33	216+3	46	.35	108+0
DOS	54/36	.35	641+4	54/41	.20	480+4	63/41	.31	565+7	64/48	.24	423+4	57/40	.31	412+2	53/39	.20	235+3

Note: *N* pairs gives first the number of complete pairs and second the number of incomplete pairs. For the DOS pair the variance of the male twin is given first, and the variance of the female twin second. Zyg = zygosity; var = variance; *r* = correlation; MZM = monozygotic male; DZM = dizygotic male; MZF = monozygotic female; DZF = dizygotic female; DOS = dizygotic opposite sex.

Discussion

The focus of this report was on the etiology of teacher-rated problem behavior at age 7, 10 and 12. Results in twin pairs rated by different teachers were compared to twin pairs rated by the same teacher. Correlations and heritability estimates were higher when one teacher rated both twins of a pair than when different teachers rated the twins. We tested if twin confusion or correlated errors could explain the difference in the correlations between the two groups. Both models gave a significant worsening of the fit of the model, thus did not provide a good explanation. This indicates an effect of GxE for boys and girls at each age. In other words, different classes with different teachers, peers and possibly classroom climate trigger different behavior depending on the child's genetic make-up. In the same teacher group, shared environment interacts with genetic risk factors, causing the twins to behave more similar in accordance with their genotype (and thus showing a higher heritability). A lower heritability is found in the different teacher group with the nonshared class environment, triggering different behavior in twins.

Our estimates of the effect of additive genetic, shared and unique environmental factors are roughly comparable with studies in which same or different teacher ratings were excluded from the analyses. Hartman et al. (2007) found a heritability of 74% for Attention deficit hyperactivity disorder (ADHD) in 106 twin pairs rated by the same teacher. Towers et al. (2000) estimated the heritability for EXT at 31% and for INT at 15% in a sample of 88 twin pairs rated by different teachers. Given the large differences between the heritabilities in the two groups, it is not surprising that our results differ from studies in which twins rated by the same and by different teachers were analyzed simultaneously (Haberstick, Schmitz, Young and Hewitt, 2005; Vierikko, Pulkkinen, Kaprio and Rose, 2006). When taking the same and different teacher ratings' ratio into account, our results are roughly comparable to the findings of Haberstick et al. (2005) in a total sample of 382 twin pairs. Possibly due to their smaller sample size, they did not report a difference in heritability between boys and girls in EXT. The estimates that Vierikko et al. (2006) reported from a sample of 1041 twin pairs were low in comparison to the present findings. This might be explained by differences in phenotypes. This study focused on EXT, which consists of rule breaking and aggressive behavior, while Vierikko et al. (2006) only considered aggression. In contrast to a meta-analysis of the environmental contributions to child and adolescent INT and EXT, no effect of shared environmental factors was detected (Burt, 2009). In the meta-analysis the effect of shared environment on EXT was estimated around 20% for both sexes, and on INT for boys at 24% and for girls at 10%. These estimates were based on parent, child and teacher ratings. Based on teacher ratings the meta-analysis reported a large dispersion in the effect of shared environmental factors. This is in line with the large dispersion in the effects of genetic factors as reported in the introduction.

To our knowledge, this is the first study investigating the twin confusion and the correlated errors model and GxE interaction in INT and EXT behavior in twins rated by same and different teachers. The twin confusion and correlated errors models have been previously applied to teacher ratings of problem behavior. For attention problems and related traits, Simonoff et al. (1998) found evidence for the twin confusion model, whereas others favored the correlated errors model (Baker et al., 2007; Derks, 2006; Derks et al., 2007; Polderman et al., 2006). It is

Table 3. *Raw estimate and proportion of variance explained by A, C and E for internalizing (INT) and externalizing (EXT) as a function of twin pairs rated by the same or by different teachers, sex and age.*

	AGE 7						AGE 10						AGE 12					
Teacher	Same teacher			Different teacher			Same teacher			Different teacher			Same teacher			Different teacher		
	A	C	E	A	C	E	A	C	E	A	C	E	A	C	E	A	C	E
							INTERNALIZING											
Boys and girls	16.56	-	7.84	7.34	-	18.15	20.52	-	9.06	7.78	-	22.37	17.06	-	8.18	10.82	-	16.73
%	68	-	32	29	-	71	69	-	31	26	-	74	68	-	32	39	-	61
(CI)	(64-71)		(29-36)	(22-35)		(65-78)	(66-72)		(28-34)	(18-33)		(67-82)	(64-71)		(29-36)	(30-48)		(52-70)
							EXTERNALIZING											
Boys	42.64	-	9.36	24.01	-	27.25	50.27	-	13.62	29.92	-	32.95	46.51	-	10.24	30.47	-	27.56
%	82	-	18	47	-	53	79	-	21	48	-	52	82	-	18	53	-	47
(CI)	(79-85)		(15-21)	(38-55)		(45-62)	(75-82)		(18-25)	(38-56)		(44-62)	(78-85)		(15-22)	(41-62)		(38-59)
Girls	24.40	-	12.25	21.62	-	17.89	30.69	-	10.18	20.80	-	24.11	28.62	-	10.37	23.52	-	18.15
%	67	-	33	55	-	45	75	-	25	48	-	52	73	-	27	56	-	44
(CI)	(61-71)		(29-39)	(47-62)		(38-53)	(71-78)		(22-29)	(38-56)		(44-62)	(69-77)		(23-31)	(46-65)-		(35-54)

Note: CI = 95% confidence interval

possible that teacher bias plays a role in behavioral problems in the domain of hyperactivity and ADHD, but not in INT and EXT problem behavior. However, these measures, especially ADHD and EXT, are significantly correlated.

Based on the current study, it is not possible to distinguish which differences between classes play a role in the GxE interaction. Earlier studies suggest that peer behavior can induce GxE in teacher rated problem behavior. The genetic component to teacher rated aggression is moderated by peer victimization (Brendgen et al., 2008). In addition, the estimated heritability of teacher rated depressive behavior is moderated by peer rejection (Brendgen et al., 2009). Teacher quality can also moderate genetic effects. Taylor, Roehrig, Hensler, Connor and Schatschneider (2010) reported for early reading that when teacher quality is low, genetic variance is restricted. However, when teacher quality is high, children's reading ability has room to be expressed according to their genetic potential. As these examples show, both teacher and classroom (peer) factors can act as environmental moderators.

It appeared that EXT problem behavior scores of twins rated by different teachers remained higher than of twins rated by the same teacher. Moreover, twins rated by different teachers also scored higher on INT problem behavior than twin pairs that stayed together, while there was only a marginal difference between these groups at age 3. This difference seemed to increase whilst the difference in EXT between the two groups remained about the same. This indicates that twin separation had a negative effect on INT and no effect on EXT behavior. Studies on the effect of twin separation at school on problem behavior are relatively scarce. Two studies reported that twin pairs at age 7, in one study rated by the mother and in the other study rated by the teacher, who were separated displayed more problems than pairs who attended the same classroom (DiLalla and Mullineaux, 2008; van Leeuwen, van den Berg, van Beijsterveldt and Boomsma, 2005). A third study only found an effect in teacher ratings of INT problems, and not in other types of problem behavior, at age 7 (Tully et al., 2004). In this third study the authors concluded that some twins (but not all) will experience significant and long-lasting INT problems due to twin separation. Our findings support this conclusion.

Conclusions

Our study demonstrated GxE in teacher rated INT and EXT problem behavior. This implies possibilities for interventions to buffer against the genetic risk of developing INT or EXT behavior during childhood. It is very well possible that the classroom environment consisting of teacher and peers is a very important factor in the expression of problem behavior and can sometimes make the difference between a child that is hard to handle at school and a child that cannot be handled and therefore needs to attend a special school. Therapeutic interventions should therefore also involve the school. Specifically applicable to twins, our study showed that separating twins at school has a negative effect on anxious depressed and withdrawn behavior based on ratings of teachers, who observe children in the school environment.It might therefore be advisable not to separate twins unless there is an urgent reason.

SUPPLEMENT TO CHAPTER 3

Table 1. *Results of model fitting for INT and EXT at age 7, 10 and 12.*

	INT age 7						INT age 10						INT age 12					
Model	-2LL	df	c	X^2	Δ df	*p*	-2LL	df	c	X^2	Δ df	*p*	-2LL	df	c	X^2	Δ df	*p*
0. SAT	40694.6	6807					37616.0	6116					25914.4	4312				
1. ACE: 4 groups*	40714.5	6817	0	19.9	10	.03	37621.7	6126	0	5.7	10	.84	25923.1	4322	0	8.7	10	.56
2. ACE: ST = DT	40856.7	6823	1	162.1	6	.00	37778.3	6132	1	156.6	6	.00	25987.9	4328	1	64.8	6	.00
3. ACE M=F, ST	40715.5	6820	1	1.0	3	.80	37626.9	6129	1	5.2	3	.16	25923.2	4325	1	0.1	3	.99
4. ACE: M=F, DT	40718.9	6823	3	3.4	3	.33	37630.5	6132	3	3.6	3	.31	25923.6	4328	3	0.4	3	.94
5. Drop C, ST	40718.9	6824	4	0.0	1	1.00	37631.8	6133	4	1.3	1	.25	25923.6	4329	4	0.0	1	1.00
6. Drop C, DT	**40718.9**	**6825**	**5**	**0.0**	**1**	**1.00**	**37631.9**	**6134**	**5**	**0.1**	**1**	**.75**	**25923.6**	**4330**	**5**	**0.0**	**1**	**1.00**
7. Correlated errors	40800.3	6819	0	105.7	12	.00	37641.7	6131	0	25.7	15	.04	25924.3	4327	0	9.9	15	.83
8. Twin confusion	40755.9	6822	0	61.3	15	.00	37718.3	6128	0	102.3	12	.00	25970.9	4324	0	56.5	12	.00
	EXT age 7						EXT age 10						EXT age 12					
0. SAT	44954.8	6914					41357.5	6218					28530.5	4359				
1. ACE: 4 groups*	44963.1	6924	0	8.3	10	.60	41371.9	6228	0	14.4	10	.16	28538.7	4369	0	8.2	10	.61
2. ACE: ST = DT	45069.8	6930	1	106.7	6	.00	41500.8	6234	1	128.9	6	.00	28607.8	4375	1	69.1	6	.00

3. ACE M=F, ST	45025.8	6927	1	62.7	3	.00	41463.8	6231	1	91.9	3	.00	28586.1	4372	1	47.4	3	.00
4. ACE: M=F, DT	44998.8	6927	1	35.7	3	.00	41406.6	6231	1	34.7	3	.00	28562.8	4372	1	24.1	3	.00
5. Drop C boys, ST	44963.4	6925	1	0.3	1	58	41373.4	6229	1	1.5	1	.22	28541.9	4370	1	3.2	1	.07
6. Drop C girls, ST	44963.9	6926	5	0.5	1	.48	41376.9	6230	5	3.5	1	.06	28542.6	4371	5	0.7	1	.40
7. Drop C boys, DT	44965.7	6927	6	1.8	1	.18	41380.5	6231	6	3.6	1	.06	28543.5	4372	6	0.9	1	.34
8. Drop C girls, DT	**44966.9**	**6928**	**7**	**1.2**	**1**	**.27**	**41380.5**	**6232**	**7**	**0.0**	**1**	**1.00**	**28544.6**	**4373**	**7**	**1.1**	**1**	**.29**
9. Correlated errors	45043.5	6929	0	88.7	15	.00	41489.0	6233	0	131.5	15	.00	28602.3	4374	0	71.8	15	.00
10. Twin confusion	45002.5	6926	0	47.7	12	.00	41425.4	6230	0	67.9	12	.00	28571.2	4371	0	40.7	12	.00

Note: The ACE model is fitted separately to same and different teacher ratings and dependent on sex, resulting in 4 different groups. The best model fit is displayed in boldface type. A = additive genetic effects; C = shared environmental effects; E = unique environmental effects; DT = different teacher; ST = same teacher; c = compared to model; M = male; F = female.

CHAPTER 4:

CHILD CARE AND PROBLEM BEHAVIOR: A STUDY OF GENE-ENVIRONMENT INTERACTION IN YOUNG DUTCH TWINS

This chapter is submitted to *Behavior Genetics* as:
Middeldorp CM, Lamb DJ, Vink JM, Bartels M, van Beijsterveldt CEM, Boomsma DI. Child care and problem behavior:a study of gene-environment interaction in young Dutch twins.

Abstract

Formal child care may influence the level of behavioral problems in children, but may also interact with genotype and environment. We assessed both effects at ages 3, 5, and 7 years for Internalizing and Externalizing problems in a large twin cohort, also considering effects of sex, and family socio-economic status.

There was a small association between formal child care and higher externalizing problems, especially when family socio-economic status was low. Environment had a larger effect in children who attended formal child care and in children from lower socio-economic status families, mainly at age 7, leading to a lower heritability for externalizing problems in these groups.

Thus, results support a bioecological model of development in which risk environments mask genetic differences between children.

Introduction

The debate whether attending formal child care outside the home during the first years of life has an effect on the development of behavioral problems in children is still ongoing. Studies have yielded contradictory results varying from negative to positive effects or no effect of formal child care on problem behavior (for an overview of studies: see Zachrisson, Dearing, Lekhal and Toppelberg (2013); Jaffee, Van Hulle and Rodger (2011); Phillips and Lowenstein (2011)). Several factors have been identified that contribute to the discrepancies in results. One review of short and long term effects of child care focused on studies performed in the United States of America (USA), including the well-known National Institute of Child Health and Human Development (NICHD) study of early child care and youth development (NICHD Early Child Care Res Network, 1998; NICHD Early Child Care Res Network, 2004; Vandell, Belsky, Burchinal, Steinberg and Vandergrift, 2010; Phillips and Lowenstein, 2011). This review reveals a fairly consistent picture of more hours spent in formal child care being associated with more problem behavior. However, this effect may be moderated by the quality of child care. For children from low income homes, spending more hours in high quality child care seems to be associated with less problem behavior (Phillips and Lowenstein, 2011). In a later USA study, not included in the review, there was no effect of child care before age 3, after correcting for family background (Jaffee, Van Hulle and Rodgers, 2011). Results from European studies are more heterogeneous (Jaffee et al., 2011; Zachrisson, Dearing, Lekhal and Toppelberg, 2013). Two early studies performed in Sweden (Andersson, 1992; Andersson, 1989) showed an association of child care and lower levels of problem behavior and this effect was larger when children entered formal child care before one year of age. Other European studies found no or a negative effect of hours spent in child care on problem behavior (Averdijk, Besemer, Eisner, Bijleveld and Ribeaud, 2011; Zachrisson et al., 2013). The relationship with socio-economic status (SES) also differs from studies based on data from the USA. In Norway, attending formal child care is more frequent in families with higher maternal education and higher family income and is associated with slightly higher scores on problem behavior (Bekkhus, Rutter, Maughan and Borge, 2011). In the Netherlands, more problem behavior as a function of child care was reported for lower SES groups (van Beijsterveldt, Hudziak and Boomsma, 2005).

Phillips and Lowenstein (2011) pointed out that, in addition to the circumstances related to child care and family background, the effect of child care might depend on neurobiological based responses to stress and on the temperament of the child. Pluess and Belsky (2009, 2010), for example, found that children with a difficult temperament, in comparison to less difficult children, experienced more behavioral problems when the quality of the formal child care was low, but fewer behavioral problems when experiencing formal child care of high quality (Pluess and Belsky, 2010; Pluess and Belsky, 2009). These effects were found at age 4.5 as well as at ages 10 and 11 years. Other studies have also indicated that a child's temperament moderates the effect of type, quantity and quality of child care, with negative effects in difficult children, but not in easy children (Crockenberg and Leerkes, 2005; Dettling, Parker, Lane, Sebanc and Gunnar, 2000).

Childhood temperament is moderately heritable (Saudino, 2005; Saudino, 2009; Groen-Blokhuis, Middeldorp, van Beijsterveldt and Boomsma, 2011). When the expression of a child's genetic susceptibility to problem behavior depends on child care, gene-environment interaction (GxE) is implied. One approach to test GxE is by the use of a twin design in which environmental exposures are assessed and the influence of genetic factors and of non-genetic factors on a trait are estimated conditional on the measured environmental exposure (Boomsma and Martin, 2002; Eaves, 1984; Purcell, 2002). If the influence of genetic factors differs as a function of exposure, this constitutes evidence for GxE (Kendler, 2001).

One twin study investigated GxE for externalizing problems (EXT) in preschool children, where preschool was defined as "formal, center-based day-care programs that include didactic learning objectives" (Tucker-Drob and Harden, 2013). At age 5, though not at age 4, EXT behavior was more heritable (67%) in children who attended preschool than in children who did not (heritability of 19%). There was no main effect of preschool attendance on EXT problems. Twin studies often distinguish non-genetic factors into environmental influences that are common to children growing up in the same family (C) and environmental influences unique to each child (E). In the same study Tucker-Drob and Harden (2012) also investigated the interaction of preschool exposure with common (CxE) and unique environment (ExE) and observed significant CxE. Common environment explained 52% of the variance in EXT problems in children who did not attend preschool, but did not explain any variance in children who went to preschool.

Two theoretical models of GxE make different predictions about the direction of the interaction. The diathesis–stress model predicts that genetic vulnerability (diathesis) in the presence of environmental stress will increase the likelihood of behavioral problems (e.g. Rende and Plomin (1992)) and also predicts that the heritability of the trait will be higher for children in at risk environments. If formal child care is seen as a risk factor that increases the likelihood of behavioral problems, the diathesis–stress model predicts a higher heritability in groups that attend formal child care.

In contrast, the bioecological model predicts that risk environments will mask genetic differences between children, whereas enriched environments will enable underlying genetic differences to be amplified (Bronfenbrenner and Ceci, 1994; Scarr and McCartney, 1983). If formal child care is seen as a risk factor, this model predicts that the heritability of the trait or the risk locus will be lower in children who attend formal child care.

We contrast these two theoretical models in a large study in twins of different ages, from a relatively homogeneous population. In addition to the effects of formal child care on problem behavior, quality of child care, genotype, age, and gender of the child, and the effects of familial SES were considered. The goal was to investigate main and interaction effects on internalizing (INT) and EXT problem behavior, where the general term 'problem behavior' is used when INT and EXT problems are considered together. Maternal ratings of problem behavior at ages 3, 5 and 7 years were analyzed in mono- (MZ) and dizygotic (DZ) twins, for whom information on formal care versus home care between ages 0 and 4 was available. Formal child care ends at age 4, because Dutch children go to school at this age. The sample consists of twins born between 1986 and 2005 who received child care between 1986 and 2009. Birth cohort was included in the model as a proxy for quality of child care with the twins divided into two groups, born before 2000 and in 2000 and after. In this period, there were several changes in the Netherlands that led to a lower quality of child care. In 1990, the national Child Care Stimulation Measure was established aiming to facilitate working mothers, which resulted in a large increase in the number of places for children in formal child care centers (from 17,000 to 59,000) (Deynoot-Schaub and Riksen-Walraven, 2005). In 1995, child care policy was decentralized to the community. The number of child care centers kept growing, but the new centers were mostly nonsubsidized as opposed to predominantly government-subsidized centers prior to 1995 (Deynoot-Schaub and Riksen-Walraven, 2005; Vermeer et al., 2008). The number of children attending formal child care rose from ~4% of children below 4 years in 1991 to 20% in 2002 (van IJzendoorn, Tavecchio and Riksen-Walraven, 2004). The increase in available places came at the expense of quality. No centers of low quality were observed in 1995 and 37% were of high quality, while in 2001, 6% of the centers scored low (36% in 2005) and only 18% scored high (0% in 2005) (Vermeer et al., 2008). The decrease in quality seems to continue. In 2008, quality was lower than in 2005 (Kruif et al., 2009). From similar data from the United Kingdom, Canada, Midwestern USA and Portugal it appears that only Portugal scored lower than the Netherlands (de Kruif et al., 2009).

Method

Subjects

Parents of twins took part in large-scale survey studies of the Young Netherlands Twin Register (YNTR), established in 1987. Data collection is longitudinal, with 2 to 3 yearly assessments and recruitment of newborn twins as an ongoing process(Bartels et al., 2007; Beijsterveldt et al., 2013). Data from the birth cohorts 1986 - 2005 were included in the current study. The surveys for the mothers included questions about formal child care when the twins were 3 and 5 years. From a total sample of 39,088 twins with data at either age, there were 25,416 twins with data both at ages 3 and 5. The response rate at age 5 for children whose mothers participated at age 3 was 77%, but no surveys at age 5 were collected in 2008 and 2009, due to a transition to a new administration database and a shortage of staff.

For 3,602 twins, the mother reported other child care arrangements than formal child care or home care, e.g. a childminder or medical day care, and for 1,892 twins, answers at the two ages were inconsistent. Data for these 5,794 twins was therefore excluded. Additional exclusion criteria were discordance for child care arrangements within a pair (N=205 pairs), a severe handicap in at least one child from a pair (N=414 twins), formal child care attendance shorter than one year (N=110 twins), or missing data on SES (N=56 twins). This left 18,932 twins, of which 3,878 had attended formal child care. The home care group also included children who went to a playgroup between ages 2.5 and 4, usually for 3 hours once or twice a week. Maternal ratings on problem behavior were available for 9,276 complete and 134 incomplete (i.e., one twin of a pair) 3-year-old twin pairs, 9,416 complete and 13 incomplete 5-year-old twin pairs and 6,218 complete and 129 incomplete 7-year-old twin pairs. The drop in the number of twins with age partly reflects the ongoing data collection and the fact that no surveys at age 7 were collected in 2008 and 2009. The response rate for age 7 in mothers who participated at age 3 was 69%. There were 1,474 MZ male, 1671 DZ male, 1719 MZ female, 1543 DZ female and 3059 dizygotic twin pairs of opposite sex (DOS). For same-sex twin pairs, zygosity was based on survey data for 5331 and on DNA markers for 1076 pairs.

Measures

For twins aged 3, the survey included questions about what type of child care the parents had arranged for their twins in the last three months. At age 5, the survey included the questions: “Did the twins go to child care outside home before age 4?” and “What kind of child care outside home did they attend?”. Answer categories were child care center, playground for toddler, medical day care center and an ‘other‘ category of nonparental child care.

Birth cohort was used as an indication of formal child care quality. The year 2000 was defined as cut-off for the switch in quality.

Family SES was based on the parent with the highest SES level. At 3 and 7 years of the offspring, a full description of the occupation of the parents was available for two-thirds of the sample and SES was coded according to the standard classification of occupations (CBS; 2001). For parents who completed an earlier version of the questionnaire, SES was based on the EPG-classification (Erikson et al., 1979), combined with information on parental education. All families were classified into three SES levels. The association between SES and behavioral problems was similar in both groups.

Problem behavior was assessed by the Achenbach System of Empirically Based Assessment (ASEBA) at 3 and 7 years of age. The Child Behavior Check List (CBCL) includes two broad categories of problem behavior: INT and EXT. Two age-appropriate versions of the CBCL were used. At age 3, parents reported on the behavior of their children by means of the CBCL 2/3 and, after 2004, the CBCL 1½ -5 (Achenbach, 1992). INT is assessed by the syndrome scales: “Anxious” and “Withdrawn/Depressed” and consists of 19 items, and EXT is assessed by: “Aggressive”, “Oppositional” and “Overactive” and includes 31 items. For twins aged 7, parents completed the CBCL 4-18 (Achenbach, 1991a) which assesses INT by the “Withdrawn”, “Somatic Complaints” and “Anxious/Depression” scales (31 items), and EXT by the “Rule Breaking” and “Aggressive” syndrome scales (33 items). For INT at age 3 and 7, Cronbach’s alphas were respectively .79 and .81. For EXT Cronbach’s alphas were .92 and .89.

At age 5, problem behavior was assessed with a subset of 42 items of the Devereux Child Behavior (DCB) rating scale. The data for subscales “Aggressive Behavior” (AGG) (7 items) and “Anxiety Problems” (ANX) (5 items) were included in the study (van Beijsterveldt, Verhulst, Molenaar and Boomsma, 2004; Spivack and Spotts, 1966). Cronbach’s alphas were .75 and .62. The DCB AGG subscale overlaps with 4 items of the CBCL EXT scale, and 2 items of the ANX DCB subscale overlap with items from the CBCL INT scales.

Statistical analyses

The independent variables (child care, birth cohort, SES, sex and SES*child care) were analyzed as predictors of a child’s problem behavior score. They were coded 0, 1 for child care (Home care versus Formal child care); 0, 1 for birth cohort (born before 2000 or born in 2000 or later); 0, 1, 2 for SES (low, middle, high) and 0 for boys and 1 for girls. The interaction of SES*child care was coded as 0 for the total group with SES 0, as 1 for formal child care with SES 1, and 2 for formal child care with SES 2.

The independent variables were also evaluated as moderators of the latent additive genetic (A), common environmental (C) and unique environmental (E) factors (see Figure 1). MZ twins share (nearly) 100% of their genetic material whereas DZ twins share on average 50% of their segregating genes. Therefore, the correlation between factor A in twins is one in MZ twin pairs, and 0.5 in DZ twin pairs. The correlation between the C and E factor in pairs is by definition 1 and 0, respectively, for all twin pairs. The effect of the latent A, C and E factors on problem behavior is given by factor loadings a, c and e. These effects can be moderated by measured variables such as child care, i.e., the effect of A can be higher or lower when a child attends formal child care. The square of the factor loadings a, c and e gives the variance explained by the genetic and non-genetic factors respectively, and these variances can change as a function of the moderator. A commonly used descriptive is the heritability of a trait, which is defined as the genetic. Note that the heritability can differ between exposed and non-exposed subjects when there is AxE or when there is CxE or ExE, because when the total variance differs between exposed and non-exposed children, the genetic variance divided by the total variance also differs.

Structural equation modelling (SEM) as implemented in the software package Mx was used to estimate parameters and test the main and interaction effects (Neale M.C. et al., 2006).

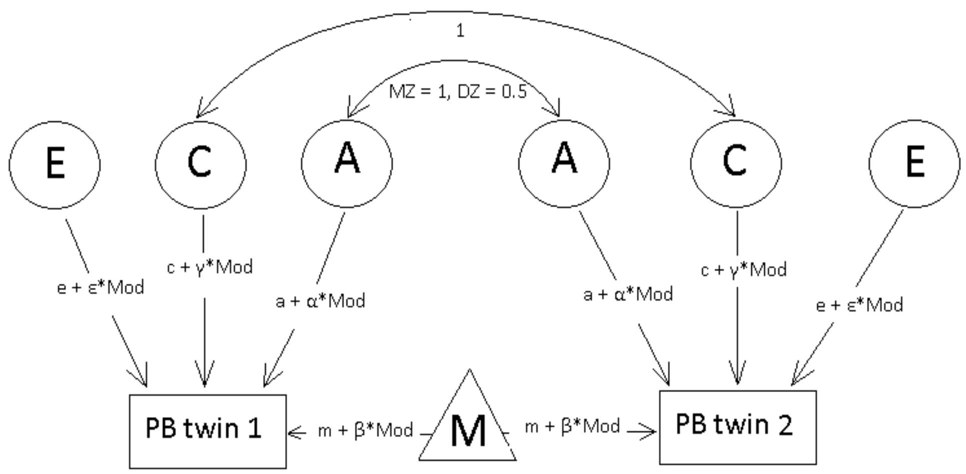


Figure 1. Gene-environment interaction model. Measured problem behavior (PB) is represented in rectangles, latent factors are symbolized by circles and represent additive genetic influences (A), common (C) and unique environment (E). Latent factors have zero means and unit variance. The influence of these factors on PB are given by factor loadings a, c and e. The factor loadings can be moderated by child care, birth cohort, SES and sex (symbolized as Mod). The effect of the moderators on the factor loadings a, c and e are symbolized as α, γ and ε, respectively. $PB_{jp} = (\alpha_i{*}Mod_i){*}A_{jp} + (\gamma_i{*}Mod_i){*}C_{jp} + (\varepsilon_i{*}Mod_i){*}E_{jp}$, where i = moderator 1 to 4 for individual j (1,2) within pair p. The variance of PB thus equals $V_{PB} = (a + \alpha_i{*}Mod_i)^2 + (c + \gamma_i{*}Mod_i)^2 + (e + \varepsilon_i{*}Mod_i)^2$. The twin pair covariance of PB depends on their genetic resemblance thus $Cov_{MZ} = (a + \alpha_i{*}Mod_i)^2 + (c + \gamma_i{*}Mod_i)^2$ and $Cov_{DZ} = ½ (a + \alpha_i{*}Mod_i)^2 + (c + \gamma_i{*}Mod_i)^2$. Mean PB is symbolized by a triangle and is moderated by the variables sex, SES, child care, birth cohort and the interaction term SES*child care, the effect of the variables on the mean is symbolized by β and equals: $m + \beta_k{*}Mod_k$, where k = 1,5.

The effects of the independent variables on the means, the factor loadings, and the effects of the independent variables on the factor loadings were estimated by maximum-likelihood. The significance of all effects was tested by constraining it at zero while other effects were included in the model. Significance of parameters was tested by a likelihood-ratio test, which indicates the decrease in likelihood for a more restricted model and which has a chi-squared distribution. When testing the main effect of child care and of SES as predictors of behavioral problems, the SES*child care interaction was omitted as main and interaction effects are not independent. In the genetic analyses, the SES*child care interaction term was not significant as a moderator of the factor loadings and is therefore excluded from the models presented in this paper. In addition, for the significant effects on the mean scores, effect sizes are given. We calculated Cohen's d (the difference between the mean divided by the standard deviation) with 0.2 to 0.3 indicating a small effect, around 0.5 a medium and larger than 0.8 a large effect.
By including all moderating variables in the means model as fixed effects when testing for GxE, we corrected for possible gene-environment correlation (rGE), taking into account the possibility that attending formal child care and levels of problem behavior might not be independent. Not considering rGE may cause biased estimates of GxE (Purcell, 2002).

Results

In line with the developments in the Netherlands, attendance of formal child care showed a steady increase between 1986 and 2005. In the birth cohorts 1986 - 1988, ~5% of the children attended formal child care. After 1989, percentages steadily rose from ~10% in the early nineties, to around 25% in the late nineties and ~30% in the beginning of the millennium. This was followed by a steep increase to ~46% for the birth cohorts 2003 - 2005. There was little variation in quantity of formal child care attendance. Eighty-three percent of the children went 2 years or more to formal child care and 65% went 3 years or more. Moreover, 76% of the parents reported that their twins went 2 or 3 days to formal child care. Therefore, the majority of the children went to formal child care for 2 years or more and attended 2 or 3 days per week.

The mean scores for INT problems were similar in boys and girls. Mean INT/ANX in boys and girls respectively was 4.6 (SD = 3.9) and 4.5 (SD = 3.9) at age 3, 10.6 (SD = 3.2) and 11.0 (SD = 3.4) at age 5 and 4.4 (SD = 4.5) and 4.7 (SD = 4.6) at age 7. For (EXT) problem behavior, mean scores were higher in boys than in girls at all ages. Mean EXT/AGG in boys and girls was 16.7 (SD = 10.0) and 14.8 (SD = 9.3) at age 3, 12.4 (SD = 3.7) and 11.7 (3.4) at age 5 and 8.5 (SD = 7.2) and 6.2 (SD = 5.8) at age 7.
Figure 2 shows mean problem behavior as a function of birth cohort, age, SES and child care and Table 2 shows the model statistics and Cohen's d. As expected, children from higher SES groups show less problem behavior. EXT/AGG was higher in children who were in formal child care at each age and the effects on EXT appeared to be largest in children from low SES families. It should be noted, given the multiple testing, that the effects were only marginally significant ($p < .05$) for EXT at age 3 and age 7, even though sample size is large. The association between child care and INT/ANX was only significant at age 5 without a significant SES*child care effect.

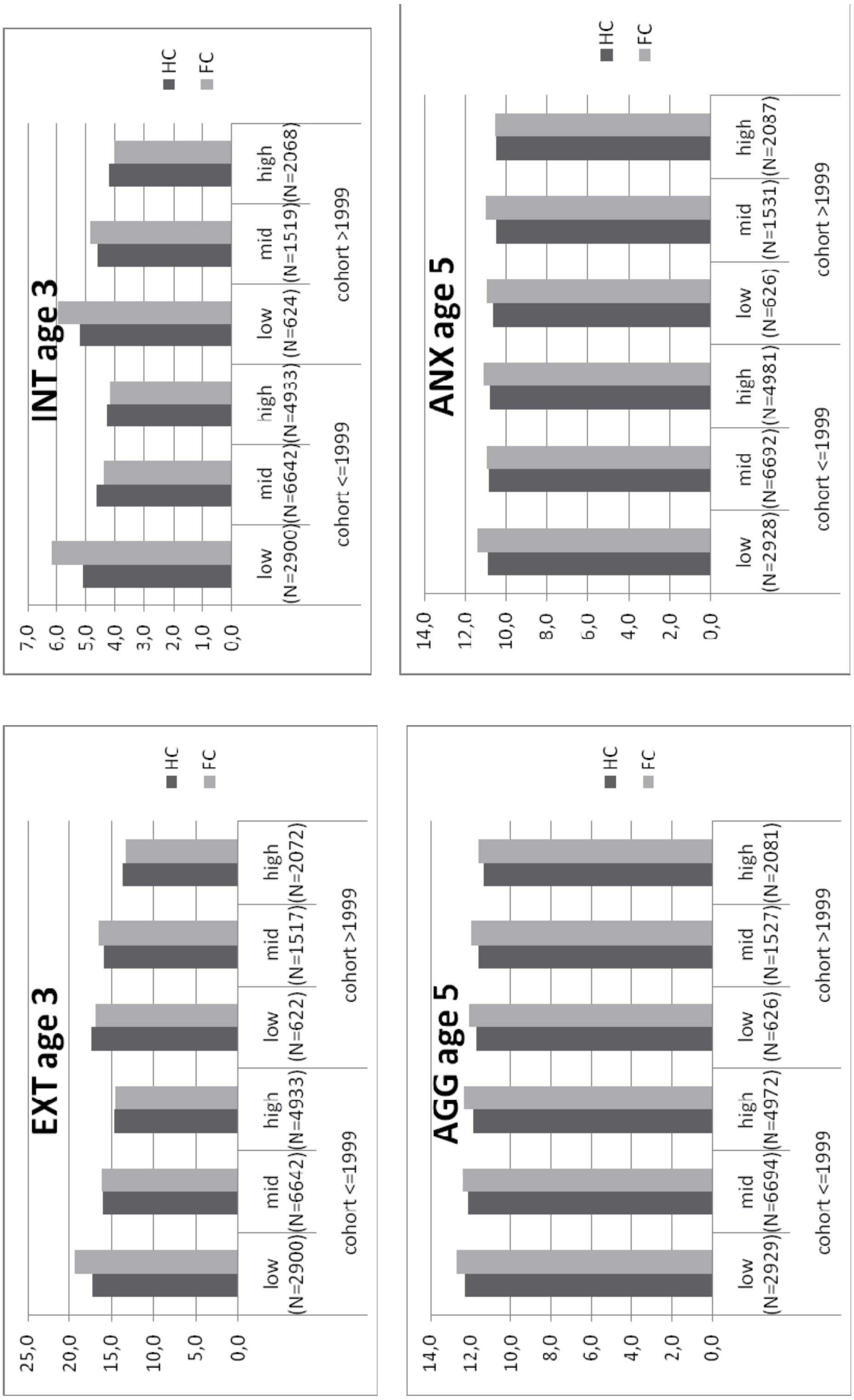
Children from later birth cohorts had lower INT and EXT scores. Moreover, Figure 2 shows that the difference between the home care and formal child care groups was not larger in the subjects born after 2000, thus the decrease in quality does not seem to be associated with more problem behavior in the formal child care compared to the home care group. Overall, the effects were small (Cohen's d between .09 and .23 for the comparison of home care and formal child care).

Turning to the results of the AxE, CxE and ExE analyses, the first half of Table 1 shows results for EXT/AGG and the second half of Table 1 gives the results for INT/ANX. First, the estimates of factor loadings a, c and e are given and, next, the effects of the moderators on a, c and e. The table shows all results with a p value < .05 so that consistencies in results over age or phenotype are not overlooked due to correction for multiple testing. However, the results with a p-value < .05 do not show such consistencies. Therefore, we focus in our presentation on the results with a p < .005. For formal versus home child care, there is significant ExE for EXT and for INT at age 7. Formal child care is associated with a larger effect of environment on problem behavior. A similar effect was seen for EXT at all ages for SES and non-genetic factors, either C or E, with higher SES associated with smaller environmental effects. For INT, the interaction with SES and non genetic factors was most significant at age 7 (p < .005), just significant (p < .05) at age 3 and non-significant at age 5. These interactions lead to a smaller total variance in the higher SES groups and consequently to a larger heritability.

To illustrate the effect of these results on the standardized variance components, Figure 3 shows the standardized variance components for EXT/AGG and INT/ANX in boys and girls born before 2000, split by age, SES group, and the home care and child care groups (effects for the groups born in 2000 and after are not depicted separately as the effects of birth cohort were, negligible). At age 3 and 5, moderation effects of SES and child care are small. At age 7, there is an increase of environmental variance with, consequently, a lower heritability in the formal child care group. In addition, the effects of E and C become smaller with higher SES leading to a higher heritability in this group. The effects are most pronounced in EXT. In boys and girls, the difference in heritability between the formal child care group of low SES and the home care group of high SES is 30% for EXT compared to a difference of ~20% in INT.

Discussion

In a large sample of Dutch twins at age 3, 5 and 7 years, we tested two competing theoretical models (i.e. the bioecological versus the diathesis -stress) regarding the question whether problem behavior is influenced by formal child care and whether the influence of genetic and environmental factors on problem behavior is moderated by the attendance of formal child care. Children who experienced formal child care scored slightly higher on EXT, especially when these children came from a lower SES family. The association between formal child care and EXT was not influenced by formal child-care quality as children born after 1999 (when quality of formal child care in the Netherlands was lower) did not Differ from children born before 2000. Overall, heritability estimates were around 50% and 60% and the common as well as the unique environment explained the remaining part of the variance. Only at age 7, the heritability of INT was lower, around 30% and 40%. Earlier studies in the NTR reported similar estimates (Bartels et al., 2004; van Beijsterveldt et al., 2004). With respect to GxE and heritability, the impact of environment was higher in the formal child care group, leading to a lower heritability. More-


INT age 3
EXT age 3
ANX age 5
AGG age 5
HC
FC
low
mid
high
cohort <=1999
cohort >1999
(N=2900)
(N=6642)
(N=4933)
(N=624)
(N=1519)
(N=2068)
(N=622)
(N=1517)
(N=2072)
(N=2928)
(N=6692)
(N=4981)
(N=626)
(N=1531)
(N=2087)
(N=2929)
(N=6694)
(N=4972)
(N=1527)
(N=2081)

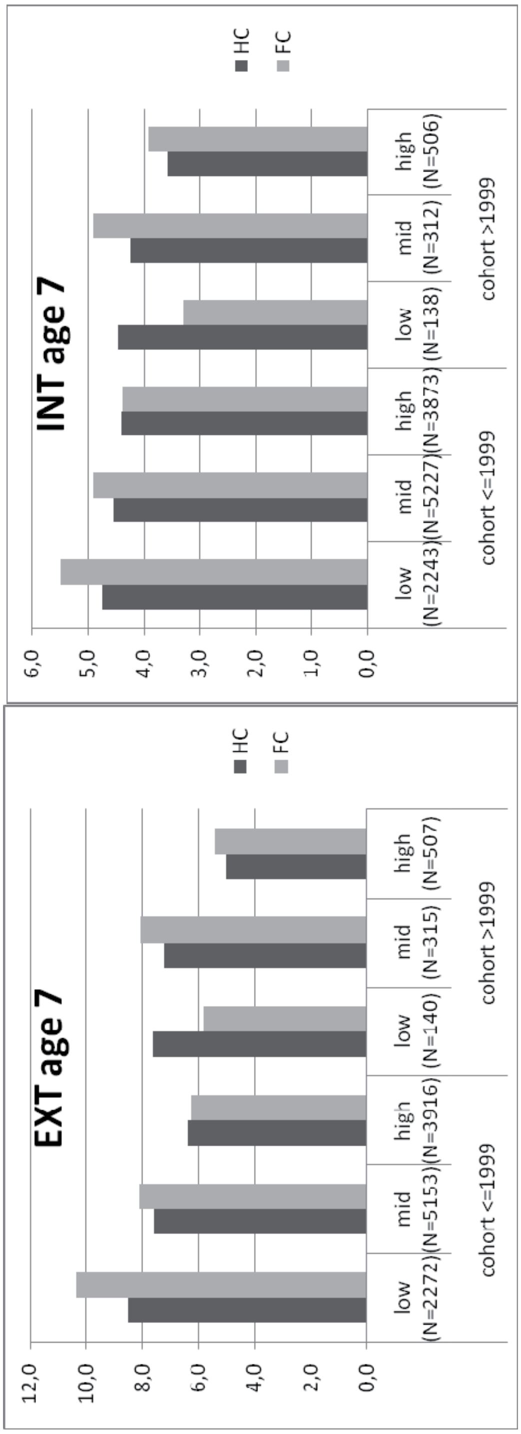


Figure 2: Mean internalizing (INT) and externalizing (EXT) problem behavior scores for home care (HC) and formal child care (FC) children born before 2000 and after from low, middle or high SES groups.

Table 1. *Parameter estimates from the gene-environment interaction model for EXT/aggression and INT/anxiety at age 3, 5 and 7. Estimates are given for factor loadings a, c and e, and for significant effects (standard errors in brackets) of the moderating variables sex, SES, child care and birth cohort on the factor loadings. The next rows give estimates for means and effect of sex, SES, child care, birth cohort, and SES*child care on the means, and effect size (Cohen's d).*

	EXT 3			**AGG5**			**EXT 7**		
	A	c	e	a	C	e	a	c	e
Factor Loading	7.64	5.76	4.41	2.71	2.01	1.96	6.36	4.07	2.75
β sex		-1.00(.35)**				-.17(.05)***	-1.30(.19)***		
β ses			-.33 (.08) ***		-.36(.12)***		-.26(.11)*	-1.31(.26)***	-.20(.05)***
β child care	-.66 (.32)*		.34(.14) *			.13(.06)*		1.12(.45)*	.45(.11)***
β birth cohort	-.73 (.29) *	.81(.35)*	-.26 (.13)*	-.45(.13)***					
Mean	18.50	Cohen's d		12.58	Cohen's d		9.73	Cohen's d	
β sex	-1.94 (.14)***	.2		-.52(.05)***	.15		-2.30(.11)***	.35	
β ses	-1.45 (.14)***	.15		-.20(.05)***	.06		-1.07(.11)***	.16	
β child care	1.39 (.57)*	.14		.30(.21)***	.09		1.53(.53)*	.23	
β birth cohort	-.58 (.22)**	.06		-.58(.08)***	.17		-.78(.26)***	.12	
β ses*child care	-.84 (.35)*	.09					-.85(.31)**	.13	

	INT 3			**ANX 5**			**INT 7**		
	A	c	e	a	C	e	a	c	e
Factor Loading	3.35	1.62	2.05	2.39	1.30	1.71	3.14	2.66	2.40
β sex		-.52(.18)*	-.14(.04)***			.13(.04)**			.16(.07)**
β ses	-.22(.07)***	-.40(.18)*	.07(.03)*					-.49(.15)***	
β child care			-4.04(.06)*				-.52(.26)*		.58(.10)***
β birth cohort		.67(.28)*	-.11(.05)*	-.32(.13)*		-.10(.06)*	-3.18(1.23)***	.98(.21)***	.37(.11)***
Mean	5.19	Cohen's d		1.75	Cohen's d		4.63	Cohen's d	
β sex	-.22(.06)***	.06		.31(.05)***	.09		.27(.08)***	.06	
β ses	-.43(.05)***	.11					-.21(.08)***	.04	
β child care				.33(.19)**	.09				
β birth cohort				-.33(.07)***	.09		-.49(.18)**	.11	
β ses*child care									

INT = internalizing; EXT = externalizing; * p < .05, ** p < .01, ***p < .005.

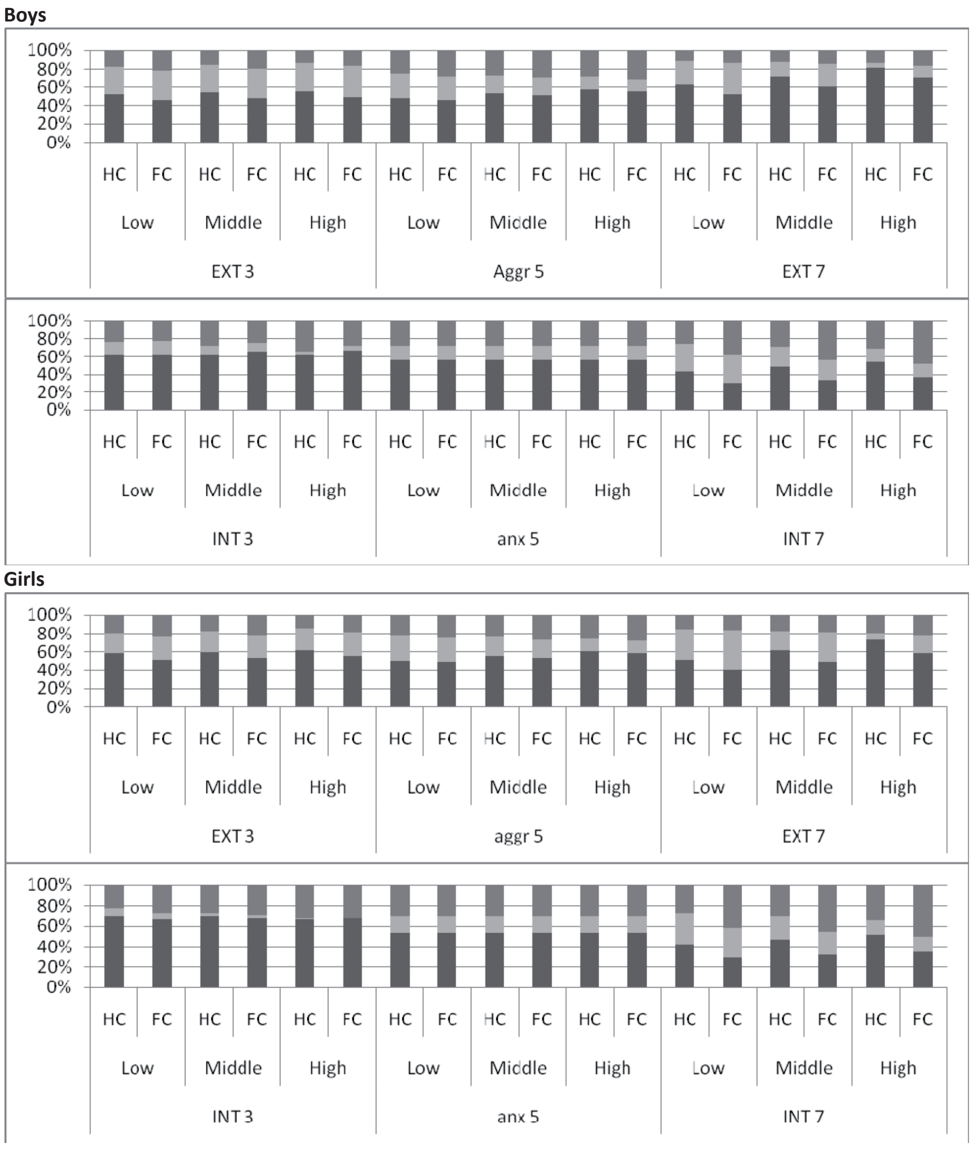


Figure 3: Standardized estimates of additive genetic (A; lower/darkest part of the bar), shared environmental (C; middle/light) and unique environmental effects (E; top/dark) for problem behavior in boys and girls with low, middle or high SES in home care (HC) or formal child care (FC).

over, children in the lower SES group were more influenced by environmental effects, thus heritability was lower in those groups. The effects of child care and SES on the influence of genetic and environmental factors were most pronounced at age 7 and for EXT.

These results are in line with the bioecological model that states that genetic-make-up becomes less important in adverse environments (Bronfenbrenner and Ceci, 1994) and not with the stress-diathesis model that predicts the opposite, i.e., that heritability is higher in adverse environments. Vendlinski, Lemery-Chalfant, Essex and Goldsmith (2011) made an overview of selected GxE studies for INT and EXT problems, including antisocial behavior. That overview shows that in five of the nine studies, results are in line with the stress diathesis model (Hicks et al., 2009; Lau and Eley, 2008; Lau, Gregory, Goldwin, Pine and Eley, 2007; Rice, Harold, Shelton and Thapar, 2006; Silberg, Rutter, Neale and Eaves, 2001), in three with the bioecological model (Button, Scourfield, Martin, Purcell and McGuffin, 2005; Hicks et al., 2009; Tuvblad, Grann and Lichtenstein, 2006) and in one with both models depending on the environmental exposure (Button, Lau, Maughan and Eley, 2008). Vendlinski et al (2011) also analyzed GxE for INT problems and several environmental risk factors and find support for both models as well. The more recent study of Tucker-Drob and Harden(2012) reports higher heritabilities(2012) report higher heritabilities and less influence of the common environment on problem behavior in children who went to pre-school. Clearly, results of GxE studies in childhood do not show a uniform picture. An obvious explanation for the discrepant findings is that there are many differences across studies, such as the age of the subjects (mean age range from 8 years till 17 years), environmental risk factors, raters (maternal or self report) and countries. Even the study of Tucker-Drob and Harden (2012) shows differences with the current study in the measured environmental risk factor. Preschool attendance rates were low before age 2 (only 15%) and increased to 72% before age 4, whereas in our sample, 83% of the children attending formal child care went there for two years or more (Tucker-Drob and Harden, 2013). Moreover, formal child care and preschool seem to be different as preschool was defined as "formal, center-based daycare programs that include didactic learning objectives". Differences in findings across age and phenotype already become apparent from our analyses in which the effects are most significant for EXT and for age 7. In sum, conclusions regarding GxE for problem behavior will depend on the environmental risk factors that are investigated and both the stress-diathesis model and the bioecological model can apply to problem behavior.

The results for the main effects on problem behavior also need to be carefully interpreted. The effect sizes were small according to Cohen's d. Some argue that the size of Cohen's d cannot be interpreted as such but should be seen in the context (Hedges, 2008; McCartney and Rosenthal, 2000). Comparing the magnitude of the effect of formal child care with the magnitude of the effect of SES on problem behavior, it is clear that the effect of SES is larger. The effect of formal child care was only marginally significant ($p < .05$) for EXT at two of the three ages despite the large sample size. Therefore, our conclusion would be that negative effects of formal child care in the Netherlands are negligible, even though the quality of care has decreased. Again, the results for the main effects underline that conclusions regarding child care cannot be easily generalized across countries (Averdijk et al., 2011; Zachrisson et al., 2013; Jaffee et al., 2011).

Overall, for the Dutch situation, both the main and the interaction effects of formal child care are small, especially in comparison to the main and interaction effects of SES. Since SES seems to have long lasting effects (e.g. Ramanathan, Balasubramanian and Krishnadas (2013)), this appears to be a more important environmental risk factor to pursue than formal child care.

A limitation of the current study is that we did not include quantity or age at formal child care entry in the analyses. However, variation in these variables was limited which precluded a useful investigation of their importance. Another issue is that attending formal child care might be associated with other variables, such as your genetic make-up (gene-environment correlation). This implies that the association between formal child care and EXT does not need to be causal, but may be explained by these other factors, as suggested by Crosby, Dowsett, Gennetian and Huston (2010) and by Jaffee et al. (2011). We controlled gene-environment correlation by including formal child care as an effect on the mean, which ensures that we did not detect spurious gene-environment interaction. It is possible that, in this way, we have missed an effect. However, the group of twins discordant for child care was too small to perform a bivariate analysis as suggested by Johnson (2007). Moreover, the question of whether these discordant twins are a representative group remains.

Conclusion

In summary, formal child care in the Netherlands is not strongly associated with problem behavior in children despite a decrease in quality. At age 7, attendance of formal child care (but especially lower SES) are associated with larger effects of E and thus lower heritabilities in line with the bioecological model. These results underline the importance of different studies across countries to investigate the effects of formal child care. For the Dutch situation, SES seems a more important risk factor than formal child care.

CHAPTER 5:

BIRTH WEIGHT IN A LARGE SERIES OF TRIPLETS

This chapter is published as:
Lamb DJ, Middeldorp CM, van Beijsterveldt CEM, Vink JM, Haak MC, Boomsma DI. (2011).
Birth weight in a large series of triplets. *BMC Pediatrics; 11:24.*

Abstract

Background: Triplets are often born premature and with a low birth weight. Because the incidence of triplet births is rare, there are relatively few studies describing triplet birth weight characteristics. Earlier studies are often characterized by small sample sizes and lack information on important background variables such as zygosity. The objective of this study is to examine factors associated with birth weight in a large, population-based sample of triplets registered with the Netherlands Twin Register (NTR).

Methods: In a sample of 1230 triplets from 410 families, the effects of assisted reproductive techniques, zygosity, birth order, gestational age, sex, maternal smoking and alcohol consumption during pregnancy on birth weight were assessed. The resemblance among triplets for birth weight was estimated as a function of zygosity. Birth weight discordance within families was studied by the pair-wise difference between triplets, expressed as a percentage of the birth weight of the heaviest child. We compare data from triplets registered with the NTR with data from population records, which include live births, stillbirths and children that have deceased within days after birth.

Results: There was no effect of assisted reproductive techniques on triplet birth weight. At gestational age 24 to 40 weeks triplets gained on average 130 grams per week; boys weighed 110 grams more than girls and triplets of smoking mothers weighted 104 grams less than children of non-smoking mothers. Monozygotic triplets had lower birth weights than dizygotic and trizygotic triplets and birth weight discordance was smaller in monozygotic triplets than in dizygotic and trizygotic triplets. The correlation in birth weight among monozygotic and dizygotic triplets was 0.42 and 0.32 respectively. In nearly two-thirds of families, the heaviest and the lightest triplet had a birth weight discordance over 15%. The NTR sample is representative for the Dutch triplet population that is still alive 28 days after birth.

Conclusion: Birth weight is an important determinant of childhood development. Triplet status, gestational age, sex, zygosity and maternal smoking affect birth weight. The combined effects amount to a difference of 364 grams between monozygotic girl triplets of smoking mothers compared to dizygotic boy triplets of non-smoking mothers of the same gestational age. Birth weight in triplets is also influenced by genetic factors, as indicated by a larger correlation in monozygotic than in di- and trizygotic triplets.

Introduction

The incidence of triplet births is rare. In the Netherlands, up to 1980, the incidence of triplet births was once per 10 thousand births. The number of triplet births increased after the introduction of assisted reproductive technologies (ART). In 1990, the number of triplet births increased to 6 per 10 thousand. From 2000 onwards, triplet birth rates decline again, mainly because of a change in policy of fertility clinics. The Central Bureau of Statistics of the Netherlands has monitored triplet birth rates at 2 per 10 thousand births since 2005 (CBS, 2010).

The prevalence of low birth weight (BW) and preterm deliveries is high in triplet births. Both low BW and prematurity are risk factors for adverse health, cognitive and behavioral outcomes later in life, (e.g. Arnoudse-Moens, Weisglas-Kuperus, van Goudoever and Oosterlaan, 2009; Bhutta, Cleves, Casey, Cradock and Anand, 2002). Among the factors that influence birth weight gestational age (GA) is the most important (Blickstein, Jacques and Keith, 2002; Luke et al., 2002; Yokoyama, Sugimoto and Ooki, 2005). Alexander et al. (1998) described how fetal growth in triplets does not follow growth curves of singletons or twins. Triplet growth is characterized by different phases. In phase A, up to 26 weeks, triplet fetal growth is comparable to that of singletons. Phase B is roughly between 26 and 30 weeks. During phase B, there is a steady decrease in triplet growth relative to singletons, up to a difference of 20%. This is hypothesized to be due to the restricted intrauterine space. During phase C, 30 to 35 weeks, there is no further decrease relative to singletons. Triplet weight during that period is about 20% less than that of singletons. These three phases are also seen in twins, though later in time and to a lesser extent. Phase D is only seen in triplets and starts from a GA of about 35 weeks. In this phase, a marked decrease in triplet weight compared to that of singletons is seen. However only 10 – 13% of the triplets reach a GA of more than 35 weeks (Blickstein et al., 2002; Min et al., 2004).

Other factors involved in triplet BW include sex, zygosity and birth order. As in singletons, boy triplets weigh more than girls (Kato and Asaka, 2002; Luke, Brown, Hediger, Misiunas and Anderson, 2006). In twins, dizygotic (DZ) twins weigh more than monozygotic (MZ) twins. This is mainly an effect of sharing a placenta. MZ twins are, compared to DZ twins, more in competition for nutrients (Loos, Derom, Vlietinck and Derom, 1998; Dube, Dodds and Armson, 2002). In triplets a similar effect is found (Chasen, Al-Kouatly, Ballabh, Skupski and Chervenak, 2002; Min et al., 2004). However, until now the effect of zygosity on BW in triplets has been based on small samples, and a distinction within DZ trios between the MZ pair and the DZ triplet has not always been made. Only a few studies specifically focused on birth order in twin and triplet pregnancies. These studies suggest that the first-born child is often the heaviest, followed by the second born child. In triplets, the third born child most often weighs the least (Orlebeke, Boomsma and Eriksson, 1993; Yokoyama et al., 2005).

Not all three children in triplet pregnancies are similarly affected with regard to BW. Inter-triplet BW discordance is thought to be a direct effect of physiological adaption to the limited uterine environment. One triplet grows at the expense of his brother or sister. Compared to twins, BW discordance in triplets is less well documented, although the phenomenon seems to be more common in triplets than in twins (Blickstein, 2002). Especially severe discordance – defined as a difference in BW of over 35% - is higher in triplets than in twins: 3.1% in twins and 9.5% in triplets (Blickstein, Jacques and Keith, 2003; Blickstein, 2004).

Maternal smoking during pregnancy is a known predictor for low BW in children (Brooke, Anderson, Bland, Peacock and Stewart, 1989; Odendaal, Steyn, Elliott and Burd, 2009). The effect of maternal alcohol consumption during pregnancy is less clear. Some studies in singletons suggested that alcohol consumption is unrelated to BW when corrected for GA (Verkerk, Vannoordzaadstra, Florey, Dejonge and Verloovevanhorick, 1993; Gibson, Baghurst and Colley, 1983). Other studies in singletons showed an effect in mothers who consume more than 100 grams or more than 5 drinks per week, as well as an interaction between alcohol consumption and smoking during pregnancy (Brooke et al., 1989; O'Callaghan, O'Callaghan, Najman, Williams and Bor, 2003). The effect of maternal smoking combined with maternal alcohol consumption on children's BW is larger than the summed effect of each separate causal agent (Brooke et al., 1989; Odendaal et al., 2009; Okah, Cai and Hoff, 2005). A study in twins found a negative effect of maternal smoking on the regression of BW on gestational age. Hence, the twins of non-smoking mothers had a more optimal development of BW (van Baal and Boomsma, 1998). As triplets are already more growth restricted compared to twins, the effects of maternal smoking and alcohol consumption could be even more detrimental. To our knowledge, no other studies have directly looked at the effect of maternal smoking and alcohol consumption during pregnancy on triplet BW.

In the past two decades, around 37% of the triplets born in the Netherlands have been registered with the Netherlands Twin Register (NTR). IIn this study we present descriptive statistics on triplet BW and analyses of the effect of sex, zygosity, birth order, GA, and maternal smoking and alcohol consumption on triplet BW are calculated as a function of zygosity to investigate the role of genetic factors on BW. Lastly, BW discordance is described. We compare characteristics of triplets registered with the NTR with data from the Netherlands Perinatal Registry (NPR, 2011). Data from the NPR consists of the total group of triplets born in the Netherlands, including the stillbirths and children that decease soon after birth.

Method

Subjects

We use the term 'triplet' to denote one of three individuals born at the same birth, and refer to a 'trio' as three triplets born at the same birth. In total, 1966 triplets from 664 families were registered with the NTR. The sample includes 642 complete trios and 22 incomplete trios. The complete trios consisted of 125 trios comprising 3 females, 187 trios consisting of 1 male and 2 females, 207 trios consisting of 2 males en 1 female, and 123 trios consisting of 3 males. The incomplete trios consisted of 18 males and 22 females. Trios were incomplete for various reasons (e.g. in young triplets: one of the triplets was deceased; in adult triplets: not all members of a trio participated).

The Adult NTR (ANTR) registers multiples who are recruited as adults and the Young NTR (YNTR) registers multiples at birth. In Figure 1 the number of triplets per birth cohort is given. Note that birth cohort 1986 marks the division between the ANTR and YNTR, triplets born after 1986 are registered with the YNTR. The oldest trios registered with the NTR were born in 1939. Data on triplet BW came from triplets born between 1970 and 2006.

To investigate the representativeness of the NTR triplet sample, we compared our sample with the Dutch triplet population regarding parity, BW, GA and age of the triplet mother when giving birth. In addition, we investigated factors that could have influenced non-response by comparing the study sample with trios who are registered with the NTR but did not participate in our surveys. The responders and non-responders were compared on age of the mother when giving birth, maternal education and population density.

Data collection

Table 2 gives a summary of data collected in triplets. Three questionnaires include items regarding pregnancy, delivery and BW. Questionnaire 1 (Q1) is completed by mothers of YNTR triplets just after registration. Q1 inquired about the pregnancy (ART, gestational age, smoking and alcohol consumption during pregnancy and mode of delivery) and characteristics of the triplet (birth date, sex, domicile, birth weight) and about characteristics of the parents. Over the years, Q1 has been collected in 323 trios. In 2008, Questionnaire 2 (Q2) was sent to mothers of all triplets born before 2006. In addition to the questions in Q1, Q2 inquired about the characteristics of the triplets up to age 2 (e.g. growth curves, health and temperament). Q2 was sent out to 535 mothers and was returned by 264 mothers. Since the data collection in 2008, all mothers of triplets who reach age 2 receive Q2.

In 2005, a questionnaire on familial twinning (Qft) was sent to all ANTR and YNTR mothers of multiples in the NTR (Hoekstra, Willemsen, van Beijsterveldt, Montgomery and Boomsma, 2008; Hoekstra et al., 2010). This survey inquired about the occurrence of multiple births within a family, mode of conception, information on delivery and parental characteristics.

The Child Behavior Check List (CBCL; Achenbach, 1991a) was sent multiple times to parents of triplets between age 3 and 12 years of the children. In addition, a short general questionnaire on parental and triplet characteristics (e.g. parental employment and religion, triplet school achievements and health) was included.

The Youth Self Report (YSR; Achenbach, 1991b; Bartels et al., 2007) was sent to YNTR

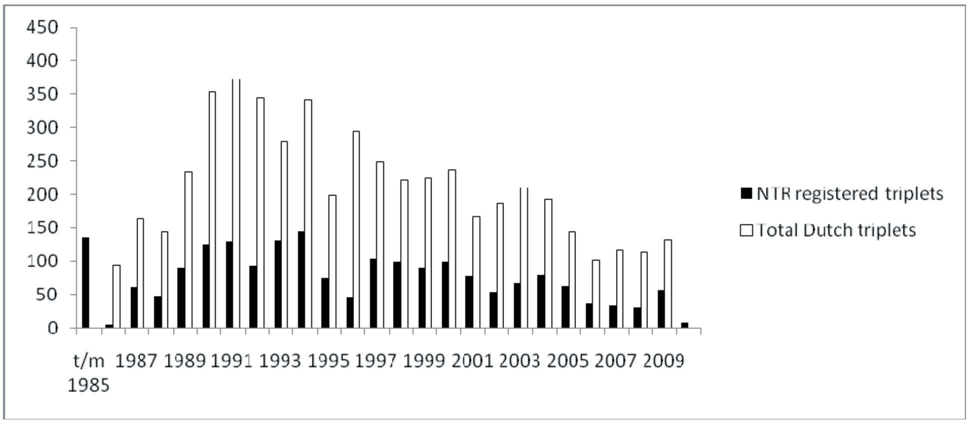


Figure 1. Number of triplets born in the Netherlands and number of triplets registered at the Netherlands Twin Register (NTR) as a function of birth cohort.

Table 1. *Descriptive statistics for a complete group of triplets born in the Netherlands between 2000 and 2006, a subset from this group including all triplets still alive after 28 days, and the NTR sample.*

	Dutch triplets (cohort 2000 – 2006)		Dutch triplets (cohort 2000 – 2006) alive after 28 days		NTR sample	
Cohort	2000 - 2006		2000 - 2006		1970 – 2006	
Parity (% primiparous)	54.7		50.9		51.9	
	M (SD)	N	M (SD)	N	M (SD)	N
BW	1745 (688)	1324	1920 (546)	1092	1900 (471)	1219
GA	32.3 (4.1)	1323	33.4 (2.8)	1092	33.7 (2.6)	1219
Age mother	31.6 (4.3)	486	31.7 (4.2)	379	30.8 (3.6)	348

triplets aged 14, 16 and 18. Triplets born between 1987 and 1992 were sent a paper and pencil version. Birth cohorts 1994 – 1995 received the YSR through the internet at ages 14 or 15. Birth cohort 1992 – 1993 completed the YSR via internet at age 16.

Starting in 2009, the Teacher Report Form (TRF; Achenbach, 1991b) is collected. In 2009 we asked parents of 170 trios for consent to approach the children's teachers. 106 parents returned the consent form and 80 gave permission. After the parental consent was obtained, teachers of 240 triplets aged 6 to 12 received the TRF.

Triplets who agreed to participate in ANTR research have usually been included in the ANTR data collection. The procedure of data collection of the ANTR is described elsewhere (Boomsma et al., 2002; Boomsma et al., 2006)

For the current study, we used the data on BW, smoking and alcohol consumption during pregnancy acquired from questionnaire Q1 and Q2, completed by the mother of the triplets. Data on ART came from Q1, Q2 or Qft. Data on BW and zygosity were available for 455 and 465 trios, respectively, out of the 642 complete trios. For 433 trios information on both was available. For 410 trios, data were available for all variables under study, i.e. BW, zygosity, GA, alcohol consumption and smoking. Information on ART was available for 329 out of these 410 trios.

Response consistency was investigated by correlating responses given at subsequent questionnaires. When comparing Q1 and Q2, correlation was 0.93 (N=521) for BW, .94 (N=182) for GA, .95 (N=182) for smoking, .57 (N=179) for alcohol consumption and .98 (N=100) for ART. Correlation for ART between Q1 and Qft was .95 (N=132), and for ART between Q2 and Qft it was 1.00 (N=132). Except for alcohol consumption, responses were highly consistent. For alcohol consumption, 14 mothers reported to have consumed alcohol on Q1 but not on Q2, and 7 mothers reported to have consumed alcohol on Q2, but not on Q1. In these cases, the response given at Q1 was used in the analysis.

Statistical analyses

Analyses were performed using the software package Mx, which allows modeling of the dependency that exists between measures of pairs of relatives (Neale M.C. et al., 2006). We tested the effect of zygosity, birth order, sex, GA, smoking and alcohol consumption on mean BW, and the effect of zygosity on variance and covariance in triplet BW. The effect of ART was tested in a smaller sub set of triplets in which information on ART was available.

First a full model, in which all effects were estimated, was fitted to the data. Subsequently, nested sub models were tested. In the full model the following parameters were estimated: the grand mean BW as a function of birth order and zygosity, the variance and covariance in BW (as a function of zygosity), and the regression coefficients on BW of GA, GA2, sex, alcohol consumption and smoking. In step 2, the means, variances and covariances of the MZ triplets from the MZ group were tested for equality with the MZ triplets from the DZ trio group. In step 3, the means, variances and covariances of the single remaining triplets in a DZ trio, not part of the MZ pair, and TZ triplets were constrained to be equal. In step 4, birth order effects on mean BW were tested within the zygosity groups. Finally, in step 5 it was tested whether the effects of GA, GA2, sex, maternal smoking and alcohol consumption on mean BW significantly differed from zero.

Using the raw likelihood method as implemented in Mx, the different models were

Table 2: *Data collection in NTR triplets. Trio list are surveys inquiring about the trio, triplet list are surveys inquiring about a triplet.*

Questionnaire	N trio list retour	N triplet list retour	Response rate
Q familial twinning	269		80%
Q1	323		58%
Q2	264		49%
CBCL age 3 – 12 mother		516	54 – 68%
CBCL age 3 – 12 father		510	54 – 68%
TRF consent form age 6 - 12	106		62%
TRF age 6 – 12		188	78%
YSR age 14 – 18		212	30 – 48%
ANTR surveys (list 1 to 8; ongoing)		315	47 – 48%

Note: CBCL = Child Behavior Check List; TRF = Teacher Report Form; YSR = Youth Self Report; NTR = Netherlands Twin Register; ANTR = Adult NTR

compared using the log-likelihood ratio test. The difference in -2 times the log-likelihood (-2LL) between two nested models has a χ^2 distribution with the degrees of freedom (df) equaling the difference in df between the two models. A p-value of 0.05 was used to determine statistical significance.

In the regression analysis, sex was coded 0 for boys and 1 for girls and GA was coded as actual GA minus 40, ranging from 0 (GA of 40 weeks) to -16 (GA of 24 weeks). A possible flattening of triplet BW at the highest GA's was modeled with GA2. GA2 was calculated by squaring the normalized score of the variable GA as described above (Mean = 0, SD = 1). Smoking and alcohol consumption during pregnancy were analyzed as dichotomous traits: yes (1) or no (0).

The MZ and DZ correlations were used to infer the influence of genetic and/or environmental factors on triplet BW. MZ triplet pairs and trios are genetically identical, whereas DZ and TZ triplet pairs and trios share on average 50% of their genetic material. A MZ correlation that is higher than the DZ correlation implies the influence of genetic factors. A DZ correlation that is higher than half the MZ correlation implies that shared environmental factors influence BW.

Birth weight discordance

Three definitions are commonly used for BW discordance in twins. First, an absolute definition: the absolute difference in BW. Second, a percentage definition: BW disparity is calculated as a percentage of the largest child's BW. Third, a statistical definition: BW as percentile of one or two standard deviations from the mean (Blickstein and Kalish, 2003). Studies on trios often adopt the percentage definition. This means that the difference between the BW of the lightest and the heaviest triplet is calculated as a percentage of the BW of the heaviest (Bagchi and Salihu, 2006).

The third method, specific to trios, takes into account the BW of the triplet that falls in between the heaviest and lightest triplets. In this method, a relative BW ratio is calculated by taking the difference between the middle and lightest triplet as a percentage of the difference between the heaviest and lightest triplet (Blickstein et al., 2003). An advantage of this method is that this ratio is representative for situations in which the trio consists of 2 heavy and 1 light triplet or trio of 1 heavy and 2 light weights.

We used the percentage definition to estimate birth weight discordance, as this definition is most frequently used.

Results

Representativeness of the NTR triplet sample

Table 1 shows parity, BW, GA and age of mother for 1) the total group of Dutch triplets born in the Netherlands between 2000 and 2006 (the NPR gathered information on birth characteristics starting from birth cohort 2000), 2) the total group of triplets but without the trios in which one or more children were deceased before 28 days after birth, 3) the NTR sample. As can be seen, the NTR sample is highly comparable to the second group, but less to the first group, which contains more primiparous mothers. In the first group, GA is shorter and BW is lower. This indicates that the NTR sample is representative of trios with a favorable outcome, i.e. children that are still alive 28 days after birth.

Comparing the current study sample with trios registered in the NTR but not participating in this survey yielded no significant age difference between the two groups ($t(454)= 1.90$, $p = .06$). There was a difference in maternal education ($\chi 2(3) = 8.69$, $p = .03$). Maternal education in the responders group was lower than the educational level of mothers from the non-responders group (percentages of low, middle, middle high and high education: 12.4%, 36.0%, 29.4% and 22.2% for the responders versus 4.2%, 41.5%, 30.5% and 23.7% for the non-responders). There was also a difference in population density between the two groups ($\chi 2(3) = 8.16$, $p = .00$). Population density was categorized as more or less than 1000 persons per square kilometre. The distribution in the response group was 50.0% and 50.0% and in the non-response group 43.4% and 56.6%, for < 1000 and >1000 persons per square meters, respectively.

Zygosity

Trio zygosity was determined by DNA, blood group assessments (vanDijk, Boomsma and Deman, 1996), or survey questions. The survey questions pertained to resemblance in hair, eye, and face color,and facial appearance of each triplet pair in a trio. Furthermore, items were included inquiring if the triplets were ever mistaken for each other by family members or strangers. When DNA, blood or survey questions were not available, self or parental report on zygosity was adopted. Self or parental report on zygosity was based on the answer to two survey questions "What do you think zygosity of the trio is?" and "And if the trio is a DZ trio, which pair forms the MZ pair?".

DNA samples were available for 79 triplets from 31 families and blood samples for 65 triplets from 22 families. Both DNA and blood samples were available for 47 triplets from 16 families. Survey questions about resemblance and self or parental report on zygosity were available for 318 and 450 trios respectively. Zygosity estimates were based on the trio. This signifies that if one pairwise comparison could be made but information on the third triplet was missing, trio zygosity could not be determined. There were 22 trios with zygosity based on DNA and/or blood information on all triplets. Seventeen trios had information on zygosity based on both DNA and survey questions regarding resemblance. This provided the opportunity to look at the reliability of the survey information. Pairwise comparisons were incorrect in 10% of the cases. This is comparable with zygosity determination based on survey questions on resemblance in twins (Rietveld et al., 2000; Willemsen, Posthuma and Boomsma, 2005). However, zygosity determination in trios is based on three pair wise comparisons. DNA and survey questions on resemblance gave the same zygosity result for 12 of these 17 trios. For 5 trios, survey questions on resemblance suggested that the trio was dizygotic while DNA determined that the trios were trizygotic. In 4 of these last 5 cases, self and/or parental report also suggested that the trios were dizygotic. Pairs determined as MZ where checked on sex (an opposite-sex pair cannot be MZ). This resulted in a zygosity determination of 465 triplet trios as presented in Table 3.

ART

ART are more commonly seen in triplet pregnancies than in twin or singleton pregnancies. We asked the 350 mothers who returned questionnaires about their possible use of ART. 127 answered that the pregnancy was spontaneous, 103 after in vitro fertilization, 17 after intracytoplasmic sperm injection, 25 after intrauterine insemination, and 63 after ovulation induction with hormone tablets or subcutaneous injections. The remaining 15 mothers gave an unclear or no answer to this question. The age of the triplet mothers who made use of ART ranged from 21 to 43 years (M = 31, SD = 3.5), the age of mothers who spontaneously conceived the trio ranged from 20 to 41 years (M = 30, SD = 3.8). ART was overrepresented in the TZ triplet group. 87% of the mothers of TZ triplets reported ART, compared to 19% of the mothers of DZ's and only 3% of the mothers of MZ triplets.

Within the subset of triplets with information on ART, we did not find a significant effect of ART on BW ($\chi2(1) = 0.23$, $p = .63$). ART was also tested in the TZ triplet group alone to correct for a possible confounding effect of zygosity, as TZ triplets are possibly heavier and overrepresented in the group of triplets born after ART. Still, no effect was found ($\chi2(1) = 1.23$, $p = .27$). All other analyses were therefore performed on the total set of triplets, including triplets without information on ART.

Birth weight

Descriptive statistics of the observed data are given in Table 4. The total sample with complete data included 37 MZ trios, 102 DZ trios which consist of one MZ pair and two DZ pairs, and 271 TZ trios. Mean GA of the triplets was little above 33 weeks. 26% of the triplets were born after a caesarean section. Only 3% of the mothers of triplets both smoked and consumed alcohol, while 15% reported smoking and 11% reported consuming alcohol during the pregnancy.

The uncorrected data on BW are presented as a function of zygosity and birth order. For the DZ trios two columns are presented. One column gives mean BW for the MZ pairs within

Table 3: *Triplet sex and zygosity*

	FFF	MFF	MMF	MMM	Total zygosity
MZ	23	-	-	23	46 (10%)
DZ	32	31	32	28	123 (26%)
TZ	31	106	125	34	296 (64%)
Total sex:	86	137	157	85	465

Note: MZ = monozygotic; DZ = dizygotic; TZ = trizygotic; FFF = trio of three females; MFF = trio of 1 male and 2 females; MMF = trio of 2 males and 1 female; MMM = trio of three males

the trio. The other column gives mean BW for the single remaining triplets that are not part of the MZ pair.

Table 4 also shows the percentage of triplets who are small for gestational age (SGA). The 10th percentile of BW as a function of GA is often classified as SGA. However, this 10th percentile differs between singletons and triplets. For example, Alexander et al. (1998) reported that for a GA of 33 weeks the singleton 10th percentile of BW is 1673 grams, but for triplets it is only 1418 grams. The discrepancy between singletons and triplets increases with increasing GA. As such data is not available for Dutch triplets, we present the percentages of triplets who are SGA based on singleton standards for the United States (US) and based on US triplet standards as reported by Alexander in addition to Swedish singleton standards which are comparable to Dutch singleton standards (Niklasson et al., 1991; Visser, Eilers, Elferink-Stinkens, Merkus and Wit, 2009).

Fit statistics of all tested models are presented in Table 5a. The effects of GA, sex and smoking on mean BW as well as the variance, covariance and correlations of BW within MZ and DZ triplets are shown in Table 5b. We found no difference between the mean, variance and co-variance of triplets from the MZ group and MZ triplets from the DZ group (step 2). There were no significant differences in the variances and covariances of the DZ and TZ triplets, but there were differences between the mean of the DZ triplets and the TZ triplets ($\chi2(3) = 16.57$, $p = .00$). A significant birth order effect was found within the TZ group ($\chi2(2) = 29.07$, $p = .00$), but not in the MZ and DZ group ($\chi2(2) = 5.61$, $p = .06$ and $\chi2(2) = 0.92$, $p = .63$, respectively).

The tests of the fixed effects showed that GA was the most important contributor to mean BW in triplets. Between a GA of 24 to 40 weeks, the triplets gained 130 grams per week. No significant flattening of the growth line (GA2) was observed ($\chi2(1) = 2.49$, $p = .11$). An effect of sex was found with boys being 110 grams heavier than girls ($\chi2(1) = 36.69$, $p = .00$). Furthermore, triplets from mothers who smoked during pregnancy were 104 grams lighter than the triplets of mothers who did not smoke ($\chi2(1) = 10.9$, $p = .00$). We did not find a significant effect of alcohol consumption during the pregnancy on triplet BW ($\chi2(1) = 0.11$, $p = .74$).

Correlations in triplet BW as a function of zygosity were calculated before and after including the effects of GA, sex and smoking on mean BW in the model. Before correction, the MZ correlation was .70, and the DZ and TZ correlations were .64 and .67, respectively. Correlations in triplet BW were lower when the effects of GA, sex and smoking were included, indicating that these variables explain part of the resemblance in triplet BW. Furthermore, the MZ triplet correlation is higher than the DZ triplet correlation, .42 compared to .32, respectively. This indicates that in addition to common environmental effects, genetic factors also explain part of the variance in BW.

Finally, BW discordance was calculated. We compared the heaviest and the lightest triplet of a trio and found that in only 17.9% of the trios, BW discordance was less than 10%. In 60.6% of the trios BW discordance was between 10 – 30% and in 21.5% BW discordance was more than 30%. A total overview of the BW discordance distribution is given in Table 6. Figure 2 presents BW discordance as a function of zygosity. There are more MZ triplets in the low discordance categories compared to the DZ and TZ triplets, and less MZ triplets compared to DZ and TZ triplets in the high discordance categories. This is in line with the higher correlation in BW in MZ triplets as reported above.

Table 4. Descriptive statistics of the raw scores on triplet BW and GA, and percentage of smoking and alcohol consumption during pregnancy, as a function of zygosity

	MZ	DZ	TZ
N trios	37	102	271
N boys / girls	41 / 70	145 / 161	387 / 426
GA mean (SD)	33.3 (2.15)	33.6 (2.89)	33.7 (2.52)
N mothers consuming alcohol	3 (8%)	19 (19%)	36 (13%)
N mothers smoking	6 (16%)	19 (19%)	49 (18%)

	MZ	MZ*	DZ**	TZ
Mean BW first born (grams)	1852	1952	2017	1954
Mean BW second born (grams)	1759	1887	1941	1933
Mean BW third born (grams)	1789	1831	2107	1829
Variance triplet BW	157	230	262	215
Pair-wise covariance triplet BW	110	162	156	144
Pair-wise correlation triplet BW	0.70	0.70	0.64	0.67
SGA first born (% SE ss/US ss/US ts)	19 / 42 / 3	22 / 38 / 5	27 / 37 / 7	25 / 44 / 7
SGA sec. born (% SE ss/US ss/US ts)	41 / 57 / 14	41 / 24 / 5	14 / 59 / 13	25 / 41 / 9
SGA third born (% SE ss/US ss/US ts)	32 / 51 / 16	35 / 39 / 8	18 / 33 / 11	18 / 52 / 15

Note * = MZ triplets from DZ trios, ** = DZ triplets from DZ trios; GA = gestational age; BW= birth weight; SGA = small for gestational age; US = United States'singleton standard; US ts = US triplet standard; SE ss Swedish singleton stnadard; MZ = monozygotic; DZ = dizygotic; trizygotic

Table 5a. *Model fitting results of the means model on triplet BW. The best fitting model is printed in bold font.*

step	Model	$N_{parameters}$	-2LL	df	to model	X^2	Δ df	*p*
1	Full model	25	576.10	1194				
2	a. Equal mean MZ pairs within MZ and DZ trios	22	577.11	1197	1	1.01	3	0.80
	b. Equal var MZ pairs within MZ and DZ trios	21	579.62	1198	2a	2.51	1	0.11
	c. Equal cov MZ pairs within MZ and DZ trios	20	581.32	1199	2b	1.70	1	0.19
3	a. Equal mean DZ/TZ	17	597.89	1202	2c	16.57	3	`0.00
	b. Equal var DZ/TZ	19	584.16	1200	2c	2.84	1	0.09
	c. Equal cov DZ/TZ	18	585.71	1201	3b	1.01	2	0.31
4	a. Equal mean within MZ	16	591.32	1203	3c	5.61	2	0.06
	b. Equal mean within DZ	14	592.24	1205	4a	0.92	2	0.63
	c. Equal mean within TZ	12	621.31	1207	4b	29.07	2	0.00
5	a. Effect $GA^2 = 0$	13	594.73	1206	4b	2.49	1	0.11
	b. Effect GA = 0	12	1013.26	1207	5a	418.53	1	0.00
	c. Effect sex = 0	12	631.42	1207	5a	36.69	1	0.00
	d. Effect alcohol = 0	**12**	**594.84**	**1207**	**5a**	**0.11**	**1**	**0.74**
	e. Effect smoking = 0	11	605.74	1208	5d	10.9	1	0.00

Note: MZ = monozygotic; DZ = dizygotic; TZ = trizygotic; GA = gestational age

Table 5b. *Parameter estimates for triplet BW and causal agents influencing BW, based on model 5d.*

	Mean/variance/covariance (grams)	
	MZ	DZ / TZ $_{(fb, sb, tb)}$
Mean	2765	2915 / 2826, 2811, 2710
Variance	111	104
covariance	46	33
correlation	0.42 (0.29 – 0.54)	0.32 (0.25 – 0.38)
covariate	Range	Effect in grams
GA	-16 to 0: weight gain per week	130
Sex	0 / 1: Boy / girl	-110
Smoking	0 / 1: No / yes	-104

Note: ***** results given for MZ triplets include triplets from both the MZ group and the DZ group; GA = gestational age, fb = first born, sb = second born, tb = third born

Table 6. Percentages of triplet pairs per BW discordance category. First presented percentage is calculated from triplets registered at the NTR, the second percentage is based on a national sample of triplets born in the Netherlands between 2000 and 2006, including triplets that deceased before 28 days after birth.

BW discordance	Heaviest - lightest	Heaviest - middle	Middle - lightest
< 10%	17.9 / 14.8 %	60.1 / 60.0 %	48.6 / 47.7 %
10 – 14.9%	18.9 / 19.1 %	18.9 / 17.2 %	18.7 / 19.1 %
15 – 19.9%	16.6 / 19.6 %	12.0 / 11.7 %	12.8 / 11.2 %
20 – 24.9%	14.1 / 12.8 %	4.9 / 5.1%	5.6 / 5.4%
25 – 29.9%	11 / 7.9 %	1.8 / 2.5 %	4.9 / 4.9 %
≥ 30%	21.5 / 25.8 %	2.3 / 3.5 %	9.5 / 11.7 %

Note: BW = birth weight

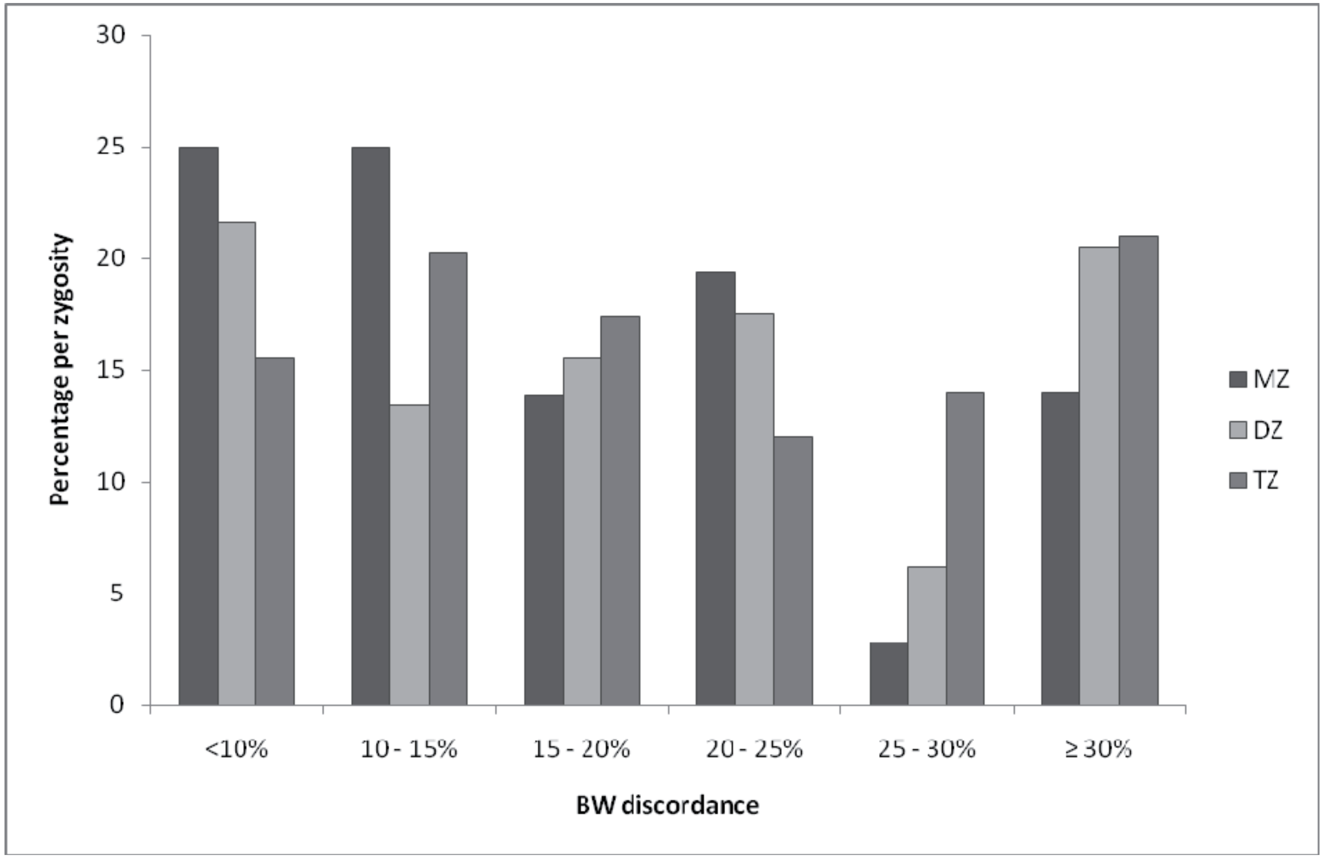


Figure 2. Percentage triplets per BW discordance category as a function of zygosity. *Note:* BW = birth weight; MZ = monozygotic; DZ = dizygotic; TZ = trizygotic

Discussion

The present study describes the influence of genetic and environmental risk factors on BW in a large population based sample of Dutch triplets. BW is affected by zygosity and birth order. MZ triplets were lighter than DZ and TZ triplets, and BW decreased with decreasing birth order in TZ triplets. GA, sex and smoking during pregnancy also had an effect on BW. No effects of ART and alcohol consumption were seen. We did not observe a significant flattening of the BW curve in the last stage of mature triplet gestation. The resemblance for BW was higher in MZ triplets than in DZ triplets after correction for the other risk factors indicating that genetic factors are also of importance. BW discordance in triplets is common since in 21.5% of the trios the difference in BW between the heaviest and lightest triplet was more than 30%.

Factors affecting triplet BW

The most important factor in triplet BW is GA. As expected, in this sample a higher GA was associated with a higher BW in triplets. The literature describes the occurrence of a flattening of the growth curve during the last weeks of triplet pregnancies. This growth restriction period emerges round a GA of about 36 to 37 weeks (Blickstein, 2004). In the NTR sample only 1.5% of triplet pregnancies reached a GA of 37 weeks or more and no flattening of the BW growth curve was seen.

The effects of alcohol consumption and smoking during pregnancy on BW were both examined. No effect of alcohol consumption on triplet BW was seen. We hypothesize that in this sample the absence of an effect was seen because of the low maternal alcohol consumption during pregnancy. In addition, mothers reported less consistent on alcohol consumption than on other variables.

Mothers of triplets who smoked had children who were on average around 100 grams lighter than non-smoking mothers, which is a decline in BW of 4%. This is the effect of smoking after correcting for the GA of the triplet. Studies in singletons report that children of smoking mothers are 119 to 241 grams lighter, which is 4 to 7% lighter than children of non-smoking mothers. The amount of loss in BW is dependent on the quantity the mother smoked during the pregnancy. Within the NTR group of smoking mothers around 70% smoked 0 to 5 cigarettes per day, 15% smoked 5 to 10, and about 15% smoked more than 10 cigarettes per day. In singletons an interaction between alcohol consumption and smoking has also been observed. We did not test for such an effect because only 3% of the triplet mothers both consumed alcohol and smoked.

First born TZ triplets were heavier than the second born TZ triplets, who were heavier than the third born children. No significant birth order effect was seen in the group of MZ and DZ triplets. As the MZ and DZ groups were small, this might reflect a lack of power to detect a difference. In twins, the first born (and heavier) twin has a heavier placenta and a more optimal (a central instead of peripheral) cord insertion (Gielen et al., 2007). Possibly triplets higher in birth order are, on average, more optimally positioned with respect to nutrients intake.

BW discordance and SGA

The prevalence of BW discordance is comparable with a study of Jacobs et al. (2003) and other studies (for a short review see Blickstein (2002) or Blickstein and Kalish (2003)). Compared to singletons or twins, triplets are delayed in growth and cognitive development. There has been limited research on the effects of BW discordance, but it seems that triplets who are discordant in BW are at an even higher developmental risk than other triplets (Feldman and Eidelman, 2005; Shinwell, Haklai and Eventov-Friedman, 2009). One study found that most triplets without BW discordance have caught up with singleton and twin standards on cognitive and executive functions at age 5. In contrast, BW discordant triplets still showed a lower performance on these functions at that age (Feldman and Eidelman, 2009). BW discordance in that study was defined as a difference in BW between the heaviest and lightest triplet of more than 15%. In the current sample, this includes 63.2% of the trios.

We also estimated the resemblance in BW of triplets. The resemblance for BW in MZ triplets was higher than in DZ triplets. Both were higher before correction for GA, sex and smoking than after adjusting for these factors. This suggests that genetic as well as common environmental factors influence birth weight and that GA, sex and smoking are some of the specific common environmental factors.

We observed that, when taking US singleton standards as a reference, 40-50% of the triplets were SGA. Children who are born SGA are at risk for asphyxia and intrauterine mortality (Gardosi, Kady, McGeown, Francis and Tonks, 2005; Pallotto and Kilbride, 2006). As a consequence, children born SGA have to be monitored in neonatal intensive care units (NICU). In the Netherlands triplets are classified SGA based on singleton standards. As a result, at least half of all triplet births have to be born in tertiary referral centers with NICU facilities.

ART

Population based triplet zygosity distributions have changed over the years. Imaizumi (2003) reported that in the Netherlands, the TZ rate increased from 1972-1973 to 1990-1991 and decreased thereafter. Imaizumi concluded that the temporary higher TZ ratio could be attributed to ART. This conclusion is confirmed by the present study in which ART was more common in TZ triplets than in the other zygosity groups.

Studies in singletons report that children born after ART are lower in BW than spontaneously conceived children. In twins the effect of ART is less clear, some studies find an effect, while others do not (Chambers, Chaptnan, Grayson, Shanahan and Sullivan, 2007; Daniel et al., 2000). The reason for lower BW in children born after ART is not completely understood. Investigators suggest that the procedure of ART itself or maternal characteristics (e.g. age, weight, parity) may cause lower BW in children born after ART. In twins, the adverse effects on mean BW associated with ART are possibly balanced by the favorable effect of DZ zygosity as ART increases the prevalence of DZ twinning and DZ twins are heavier than MZ twins. In the present study we found no effect of ART. The effect was neither present in the entire triplet group, nor in the TZ triplet group. Therefore, in present study it can be concluded that the presumed lowering effect of ART on BW was not counterweighted by the higher prevalence of TZ triplets in the ART group.

Limitations and strengths

The present sample consisted of triplets who were registered at the NTR and whose parents were willing to participate in survey studies. This led to a small positive selection bias. Triplets from families in which all three children are alive 28 days after birth also have more favorable scores on BW and GA. Moreover, parents are possibly more willing to participate in research when the triplets are healthy compared to parents dealing with illnesses of one or more of their children. The NTR sample was more comparable with a selection of Dutch triplets that were still alive 28 after birth, than with a complete group of Dutch triplets including children who died soon after birth. A comparable positive selection bias was found in a study on secular trends in gestational age and birth weights in twins. In this study twins registered at the NTR were compared with a national reference set. The twins registered at the NTR were found to have a higher GA (36.5 (2.4) compared to 35.9 (3.0) weeks) and a higher BW (2498 (550) grams compared to 2459 (615) grams)(Gielen et al., 2010). As a result of this positive selection bias, percentages of discordant triplets are probably underestimated compared to the total Dutch triplet population. The positive selection bias could also cause an underestimation of the percentage triplets that are classified as SGA. We also do not know whether the effects of the investigated risk factors might be more pronounced in this more vulnerable group.

We investigated whether zygosity influenced triplet BW. MZ triplets are more in competition for nutrients than DZ and TZ triplets. A more direct measure of triplets sharing placentas and therefore triplet nutrients competition is chorionicity. Information on chorionicity would therefore have been a valuable addition to the information on zygosity. Currently, no reliable information on chorionicity was available.

Parity has been associated with BW but was not included in the analysis of present study, as information on parity was only available for about three-quarter of the mothers. Including parity would therefore have decreased sample size considerably. In an analysis within the reduced sample, there was no significant effect parity on BW.

Some strengths of this study are also noteworthy. The sample is relatively large. We do not know of another study that took so many risk factors into account analyzing their effect on BW. This study is the first to describe triplet zygosity in the Netherlands based on individual zygosity measures instead of population based estimated zygosity distribution. We therefore could confirm the assumption that ART has inflated the Dutch TZ triplet population. Furthermore, this study is the first to investigate the effect of maternal smoking and alcohol consumption during pregnancy on triplet BW. In addition, although our sample is somewhat positively biased when comparing it to all triplets born in the Netherlands, it is a representative sample for the Dutch triplet population that is still alive one month after birth.

Conclusions

Longitudinal data collection on triplets is scarce. Data collection within the Netherlands Twin Register (NTR) is broad, including an important focus on behavior. The data collection in triplets that we are currently establishing is unique in its kind. With this dataset it is possible to study long term effects of low BW in triplets, both on physiologic and also on behavioral level. This study was limited to a description of the sex and zygosity distribution of the triplets and the effect of a number of BW characteristics. We found an effect of GA, sex, birth order, zygosity and maternal smoking on triplet birth weight, but found no effect of ART and maternal alcohol consumption. The combined effects implied that differences of 364 grams can be observed between MZ girl triplets of smoking mothers compared to TZ boy triplets of non-smoking mothers of the same GA. Furthermore, we found that MZ triplets resembled each other more than DZ triplets, indicating that, in addition to environmental factors, genetic factors contribute to triplet BW.

CHAPTER 6:

EFFECTS OF CHORIONICITY AND ZYGOSITY ON TRIPLET BIRTH WEIGHT

This chapter is published as:
Lamb DJ, Vink JM, Middeldorp CM, van Beijsterveldt CEM, Haak MC, Overbeek LIH, Boomsma DI. (2012). Effects of chorionicity and zygosity on triplet birth weight. *Twin Research and Human Genetics; 15 (2)*; 149-157.

Abstract

Birth weight in triplets is on average lower than in singletons and twins, and more children are classified as having very low or extremely low birth weight. Still, there is limited research on factors that affect triplet birth weight, and samples under study are often small. Chorionicity and zygosity influence triplet birth weight but it is unknown whether the effect of zygosity can be entirely ascribed to the effect of chorionicity or whether zygosity has an additional effect on triplet birth weight.

This question was investigated in 346 triplets (from 116 trios) registered with the Netherlands Twin Register (NTR) for whom data on chorionicity were available. 'Triplet' refers to one child and the set of three triplets is referred to as 'trio'. Trios and triplets were classified based on zygosity and chorionicity. With regression analysis the effects of zygosity and chorionicity on triplet birth weight were examined, while controlling for gestational age, sex and maternal smoking during pregnancy. In addition, within the dizygotic trios a within-family comparison was made between the birth weight of the triplets that were part of a monozygotic pair (with some pairs sharing a chorion), and the birth weight of the dizygotic triplet.

Monozygotic monochorionic triplets had a lower mean birth weight than dizygotic dichorionic triplets, based on the classification on individual level. Most remarkably, in dizygotic trios, monozygotic pairs only had a lower mean birth weight than their dizygotic sibling triplet when the pair shared a chorion. We conclude that not being MZ but rather having shared a chorion increases the risk of a low birth weight.

Introduction

Triplets often have a low birth weight. This is partly due to low gestational age; they are often born premature. However, zygosity and chorionicity also affect birth weight. Triplets, which refers to one child and the set of three triplets is referred to as a 'trio', can be either monochorionic (MC), dichorionic (DC) or trichorionic (TC). A smaller number of choria is associated with lower birth weight as sharing a placenta causes competition for nutrients (Adegbite, Ward and Bajoria, 2005; Bajoria, Ward and Adegbite, 2006; Geipel et al., 2005). In addition, there is some evidence that when more triplets share a chorion, gestational age of the trio is lower (Spencer et al., 2009).

Zygosity also is an important factor in birth weight, but is not independent from chorionicity. Monozygotic (MZ) triplets can share one chorion, can be DC, or have each their own chorion (i.e. trichorionic: TC); dizygotic (DZ) triplets are DC or TC; trizygotic (TZ) triplets are always TC (Fox and Sebire, 1997). In an earlier study we found that on average, MZ triplets have a lower birth weight than DZ and TZ triplets (Lamb et al., 2011). However, other studies did not find a significant effect of zygosity on triplet birth weight. This could be due to the size of the MZ samples (Newman, Jones and Miller, 1991; Orlebeke et al., 1993).

Studies combining information on chorionicity, zygosity and birth weight in triplets are rare. In one study describing 15 spontaneously conceived trios, no conclusions regarding the effects of chorionicity and zygosity could be made due to the small sample size (Machin and Bamforth, 1996). This study suggested that compared with the triplets born after assisted reproductive techniques (ART), being spontaneously conceived is associated with a higher proportion of MZ embryos. Many MZ embryos are MC with vascular anastomoses (i.e. connected blood vessels between the circulation of both (or more) embryos), which can result in twin-to-twin transfusion syndrome, resulting in lower birth weight and even fetal death in some cases. A second study with data on both zygosity and chorionicity, showed that based on zygosity, MZ triplets are lighter than the DZs and DZs are lighter than TZ triplets. Based on chorionicity, the DC triplets appeared to be the lightest. Zygosity and chorionicity were not combined in this study to predict birth weight and because of the low sample size none of the effects were significant (Guilherme et al., 2008).

The effect of chorionicity and zygosity on BW has been examined in twins (Dube et al., 2002; Loos et al., 1998; Loos, Derom, Derom and Vlietinck, 2005; Lee, Oh, Lee, Kim and Jun, 2010). Twins can be distinguished into three groups: MZ MC, MZ DC and DZ DC. Two studies with respectively 1080 and 569 twin pairs reported that MZ DC twins did not differ in birth weight with DZ DC. MZ MC had lower mean birth weight than the DZ DC (Dube et al., 2002; Lee et al., 2010). In contrast, a third study in 4735 twins showed that mean birth weight was highest for the DZ DC twins, lower for the MZ DC twins, but even lower for the MZ MC twins (Loos et al., 1998). In a follow up on this third study, within 4529 pairs with information on placenta fusion and umbilical cord insertion it was shown that DC twins weighted more than both MZ DC and MZ MC twins, but only when the placentas were fused and the cord insertion was peripheral (Loos et al., 2005). Otherwise, the mean birth weight for all three types of twins was similar. In general it seems that MZ MC twins have lower birth weight than MZ DC and DZ DC twins, suggesting that chorionicity is more important than zygosity. In addition, when comparing the MZ DC and the DZ DC twins, placenta fusion and cord insertion seems less optimal with respect to nutrients intake for the MZ than for the DZ suggesting a disadvantage for birth weight of MZ compared to DZ twins.

Birth weight discordance is also expected to be influenced by chorionicity as well as zygosity. MC twins were found to be more discordant than the DC (Blickstein and Keith, 2004; Gielen et al., 2008; González-Quintero et al., 2003). Gielen, Derom, Derom, Vlietinck and Zeegers (2009) hypothesized that MZ DC twins are least discordant because they do not share a placenta but are genetically identical, whereas DZ DC and MZ MC twins are assumed to be more discordant because the first share on average 50% of their genetic material and the latter are in competition for nutrients.

Based on the limited literature, it is expected that zygosity and chorionicity are independently associated with triplet BW. In this paper, we explore the effects of chorionicity and zygosity on triplet BW using information on 346 triplets registered with the Netherlands Twin Register (NTR). To obtain information on chorionicity, data were linked to PALGA, the nationwide network and registry of histo- and cytopathology in the Netherlands (Casparie et al., 2007). PALGA was founded in 1971 and has nationwide coverage since 1991. The registry was established by a number of Dutch pathology laboratories that aimed to promote information exchange between laboratories in addition to providing useful data to others in the health care field and research. Triplet placentas are sent for histopathological evaluation, including both gross and microscopic evaluation and calling of chorion type, and those data are added to PALGA.

By linking the triplet information from NTR to chorionicity information from PALGA, we aim to assess whether zygosity has an additional effect on triplet BW when chorionicity is accounted for. To investigate this question, trios and individual triplets were classified based on zygosity and chorionicity. With regression analysis, it was investigated if triplet birth weight is predicted by chorionicity and zygosity, while controlling for other factors which are known to affect triplet birth weight (i.e. gestational age, sex, maternal smoking during the pregnancy). Triplet birth weight was also classified as normal, low, very low and extremely low and percentages of each birth weight classification were given as a function of triplet and trio zygosity and chorionicity. In addition, groups based on triplet and trio zygosity and chorionicity were compared on birth weight discordance and on gestational age, as it had been suggested that with less choria in a triplet pregnancy, gestational age is lower (Spencer et al., 2009). Finally, within the DZ trios, which provided a within family comparison between a MZ pair and a DZ triplet, the effect of zygosity and chorionicity was examined.

Figure 1. Ultrasound picture at 12 weeks' gestational age of a monochorionic trio. Therefore this trio is monozygotic. Arrow indicates the meeting pointing point of three amniotic membranes. No lambda sign is visible at the side where the membranes insert the placenta. Numbers indicate the three fetuses.

Figure 2. Ultrasound picture at 13 weeks' gestational age of a dichorionic triamniotic trio. The arrowhead indicates the lambda-sign, which is prove that this fetus does not share it's placenta with fetus 2 and 3. The arrow indicates the amniotic membranes of fetus 2 and 3, which is a monozygotic pair. At this time it is unsure if fetus 1 shares zygosity with 2 and 3. Note the difference in thickness of the septum between 1 and 2/3 versus the thickness between 2 and 3. Numbers indicate the three fetuses.

Figure 3. Ultrasound picture at 12 weeks' gestational age of a trichorionic trio. The arrowheads indicate lambda-signs between each fetus. These 3 fetuses do not share their placentas. The membranes between the fetuses are thick. This trio can be trizygotic, dizygotic (one identical duo), or monozygotic. Numbers indicate the three fetuses.

Method

Participants

Triplets were registered at birth with the Netherlands Twin Register (NTR). The NTR was established in 1987 and collects longitudinal data on multiples by mailed surveys and through interviews, home visits and deep phenotyping (Boomsma et al., 2002; Willemsen et al., 2010). Questionnaire 1 (Q1) was sent to mothers of triplets soon after registration. In 2005, a questionnaire on familial twinning (Qft) was sent to all mothers of multiples, including all triplet mothers (Hoekstra et al., 2008). In 2008, all mothers of triplets received the second NTR survey, questionnaire 2 (Q2). Q1 and Q2 included items about the triplet pregnancy, birth and additional triplet characteristics (Lamb et al., 2011).

In Q2 and Qft mothers were asked for permission to link their data to other registries. For 334 mothers of triplets who gave permission to link data (88% of the mothers who returned Q2 and/or Qft), information from PALGA was requested. Data were linked on last name of the mother (maiden name), date of birth of the mother and first initial (if available). For a subgroup, a second attempt was done with birth date of the triplets as extra identifier, resulting into a total of 234 hits. Of the 100 trios without successful linkage, 23 were born before 1990, 23 maiden names of the mothers were missing and for 26 mothers we were uncertain about their maiden names. In the end, PALGA provided placenta pathology information for 234 sets of triplets (successful linkage of 70%), with the first record dating from 1978 and the last dating from 2006. For 176 trios, information about triplet chorionicity was available (75% of the sample provided by PALGA).

Triplet zygosity in the NTR data was based on DNA for 2.6% of the sample, on chorionicity for 9%, on blood group assessments for 2%, on a series of survey questions for 70.4% and on a one question survey item on zygosity for 14.5%. Triplets with inconsistent data on zygosity in the NTR and chorionicity form PALGA were excluded (N=19 trios). The zygosity of 2 trios was DZ and TZ (this was confirmed with the presence of an opposite-sex pair within the trio), while PALGA data on chorionicity indicated MC and for 17 trios information on zygosity indicated TZ while chorionicity indicated DC. In addition, zygosity was changed from DZ and TZ into MZ (N trios = 2) if only 1 chorion was present according to PALGA information.

Information on triplet birth weight, zygosity and chorionicity was available for 114 complete and 2 incomplete trios. All of these trios were born between 1987 and 2005. Age of the mother at the birth of the triplets ranged between 21 and 41 years old (M = 30.6, SD = 3.7). To investigate the representativeness of the sample, the NTR trios were compared regarding parity, birth weight, gestational age and the age of the mother at the time the triplets were born with the total Dutch triplet population based on information from the Netherlands Perinatal Registry (PRN-foundation, 2011). In addition, the analyzed sample was compared to the group without chorionicity information or without link to PALGA, on maternal education, age of the mother at birth of the trio, population density of the area the triplets resided.

Measures

Data on triplet birth weight, gestational age and on parity were obtained from two NTR surveys, Q1 and Q2. Information on ART came from the NTR surveys Q1, Q2 and Qft (van Beijsterveldt et al., 2008). ART included in vitro fertilization, intracytoplasmic sperm injection or ovulation induction with hormone tablets or subcutaneous injections.

Information on maternal smoking (yes or no) during the triplet pregnancy was obtained from Q1 and Q2. Around 90% of the mothers reported that they did not smoke during pregnancy, around 7% smoked 0 to 5 and about 2% smoked 5 to 10 cigarettes per day.

Educational level of the mother was assessed in Q2 with four categories: lower education, secondary education, upper secondary education, and higher education including academic education. As just a small percentage of mothers had received only lower education, the first and second category were merged.

Postal code of the families were linked to information of Statistics Netherlands to obtain information regarding population density for the area where the family resided (Statistics Netherlands: (2010). Population density was classified as more than or less than 1000 persons per square meter.

Data analyses

Classification Triplet birth weight was classified as normal (NBW) if birth weight was more than 2500 gram, low (LBW) if birth weight was between 1500 gram and 2500 gram, very low (VLBW) if birth weight was between 1000 gram and 1500 gram and extremely low (ELBW) if birth weight was less than 1000 gram.

Two classifications of the sample were made: (1) classification based on *trio* zygosity and chorionicity and (2) classification based on *triplet* zygosity and chorionicity. The first classification defines the trios with respect to zygosity (MZ, DZ or TZ) and chorionicity (MC, DC or TC) resulting in the following six groups: MZ MC, MZ DC, MZ TC, DZ DC, DZ TC and TZ TC. The second classification defines the zygosity (MZ or DZ) and chorionicity (MC or DC) of two children within a trio (e.g. child 1 with child 2 and with child 3). When a triplet forms a MZ pair with at least one other child, this pair is classified as MZ. In addition, when a chorion is also shared, this pair is classified as MC, otherwise as DC. When a child has a DZ relation with both other triplets, this individual is classified as DZ, who is by definition DC. This classification results in three groups: MZ MC, MZ DC and DZ DC, Table 1 gives an overview.

Birth weight discordance for *trios* was calculated as the difference between the heaviest and lightest triplet, as a percentage of the heaviest of the trio. This method is often used to quantify triplet birth weight discordance (Bagchi and Salihu, 2006; Blickstein, 2002). Next, birth weight discordance for the *triplet* classification was calculated as the discordance within the trio (i.e. as described above) for the MZ MC, MZ TC and TZ TC trios. Because not all triplets in the DZ DC and DZ TC trios received the same triplet zygosity and chorionicity classification, their birth weight discordance was calculated as follows. Birth weight discordance was calculated within the pair for the triplets forming a MZ pair. For the triplet that formed a DZ relation with both others in the trio, birth weight discordance was calculated between birth weight of this triplet and mean birth weight of the MZ pair.

Statistical analyses

The effect of chorionicity and zygosity on triplet birth weight

To examine the effect of *trio* zygosity and chorionicity on triplet birth weight, while correcting

Table 1. *Description of classification based on zygosity and chorionicity.*

Triplet zygosity	MZ	DZ	TZ
Chorionicity			
MC	N trios = 11 MZ MC, N triplets = 33	-	-
DC	N trios = 5 *Excluded **	N trios = 24 Mz pair = MZ MC, N triplets = 48 DZ triplet = DZ DC, N triplets = 24	-
TC	N trios = 2 MZ DC, N triplets = 5	N trios = 18 Mz pair = MZ DC, N triplets = 36 DZ triplet = DZ DC, N triplets = 18	N trios = 56 DZ DC, N triplets = 167

MZ = monozygotic, DZ = dizygotic, TZ = trizygotic, MC = monochorionic, DC = dichorionic, TC = trichorionic. N.B. For each type of trio, the number of trios is given (N trios) in addition to the classification of the triplets in the trio (e.g. DZ DC trios (N trios=24) can be classified into triplets that together form a MZ pair, i.e. MZ MC triplets (N individuals/triplets=48) and one triplet that has a DZ relation with both other members of the trio, i.e. DZ triplet (N triplets = 24)). *MZ DC trios were excluded because it is not possible to decide which two triplets share a chorion.

for the effect of gestational age, sex and smoking, a multiple regression analysis was performed. The variables were specified as follows: gestational age as a continuous with the actual gestational age minus 40 (range: -13 to -3), with sex (0 for boys and 1 for girls) and maternal smoking (0 for non smoking and 1 for smoking) as dichotomous and with chorionicity and zygosity as two categorical predictors with both three levels, using dummy coding for these last two predictors.

The effects of *triplet* zygosity and chorionicity on birth weight were analyzed with multiple regression analysis with gestational age as a continuous, sex and smoking as dichotomous and triplet zygosity and chorionicity as a categorical predictor with three levels (i.e. MZ MC, MZ DC and DZ DC, using dummy coding). All regression analyses were performed in Stata with correction for within cluster (i.e. family) correlation (StataCorp, 2005).

Comparing gestational age and birth weight discordance between the groups

First, with an Analysis of Variance (ANOVA) it was examined if gestational age and birth weight discordance differed between the groups classified based on trio and based on triplet zygosity and chorionicity. ANOVA was carried out when the dependent variable (i.e. gestational age, birth weight discordance) was the same for all triplets in a trio, so analyses could be performed at a family level in SPSS (SPSS Inc., 2008).

DZ DC and DZ TC trios, comparing the MZ pair with the DZ triplet

Two regression analyses were performed. First, a multiple regression with gestational age as continuous predictor and sex, smoking, being part of the MZ pair and sharing chorion as dichotomous predictors. Second, an interaction term between being part of the MZ pair and sharing a placenta was calculated, resulting in “1” for the MZ pairs that shared the chorion and “0” for all others, that is, the DZ triplets and the MZ pairs that did not share a chorion.

Results

Representativeness of the NTR sample with PALGA information on chorionicity

Information on chorionicity in the NTR sample was compared to two representative Dutch samples from the PRN-foundation (Table 2). The PRN-foundation gathered information on birth characteristics in the Netherlands as of 2000. The first sample contains data on triplet birth weight, gestational age, parity and age of the mother at the triplet birth for all triplets born in the Netherlands between 2000 and 2006, including the stillbirths and triplets who deceased soon after birth. The second sample is a subsample of the first sample and consists of triplets that were still alive 28 days after birth. It becomes clear from Table 2 that the NTR sample with data available on chorionicity is comparable to the representative Dutch triplet sample that is still alive 28 days after birth regarding birth weight, gestational age, parity and age of the mother at birth of the triplets.

The current study sample was also compared to data from the NTR triplet sample whose mothers who gave permission for retrieving information from national registries but who were not linked to PALGA, and to data from the group of triplets of whom PALGA did provide some information, but not on chorionicity (Table 3). There are no important differences between the groups that could lead to a bias in the analyzed group. Among the mothers whose data were not linked to PALGA, a somewhat higher percentage gave birth to their triplets before 1991 (25.6% versus 17.1%), which is explained by from the fact that PALGA did not have national coverage before 1991.

Normal, low, very low and extremely low triplet birth weight

The largest part of the sample had a LBW (NBW=6.4%, LBW = 71.7%, VLBW=18.8% and ELBW=3.2%). Table 4 presents an overview of the percentages of triplets according to their birth weight classification, as a function of the trio and triplet zygosity and chorionicity groups. First, as can be seen in Table 4, percentages presented for the MZ TC group should be viewed with care because these are based on small numbers. Due to these small sample sizes, no statistical testing was performed. Considering the groups classified on *trio* zygosity and chorionicity, it seemed that VLBW is more frequently present in the MC and DC groups than in the TC groups. Based on *triplet* zygosity and chorionicity, VLBW triplets are more prevalent in the MZ MC groups than the other two groups.

The effect of chorionicity and zygosity on triplet birth weight

Triplet birth weight as a function of *trio* chorionicity and zygosity is presented in Table 5. Differences between the groups could indicate an effect of chorionicity as well as zygosity. But neither chorionicity nor zygosity was significantly associated with birth weight. In Table 6, triplet birth weight is presented as a function of *triplet* zygosity and chorionicity. MZ MC triplets have a significantly lower mean BW than the DZ DC triplets (β = 86.4, SE = 0.07, t(110) = 0.047).

Comparing gestational age and birth weight discordance between the groups

Differences in gestational age and birth weight discordance between the groups based on *trio* zygosity and chorionicity were not significant (F(5,110) = 1.39, p = .23 and F(5,105) = 0.85, p = .52, respectively, see Table 5). In addition, differences in gestational age and birth weight discordance were again not significant, when groups were classified on triplet zygosity and chorionicity (F(2,108)=0.81, p = .45 and F(2,104) = 0.71, p = .50, respectively, see Table 6).

DZ DC and DZ TC trios, comparing the MZ pair with the DZ triplet

DZ DC and DZ TC trios provided the opportunity to perform a within family comparison between a MZ pair and a DZ triplet. In DZ DC trios the MZ pair shares one chorion, in DZ TC trios the MZ pair does not share a chorion. In Table 7 the mean birth weight of MZ pairs and DZ triplets are presented as a function of chorionicity. The table suggests that mean birth weight of the MZ pair was lower only when the MZ pair shared a chorion. When the two triplets of the MZ pair each had their own chorion, there was no difference in birth weight between the MZ pair and DZ triplet. After testing, neither the main effect of being part of an MZ pair within the DZ trio nor the main effect of sharing a placenta was significant in this group (β = -124.8, SE = 0.07, t(40) = -1.68, p = .10 and β = 138.8, SE = 0.01, t(115) = 0.012, p = .10 respectively). However, the interaction between being part of the MZ pair and having shared a placenta was significant (β = -131.9, SE = 0.06, t(40) = -2.39, p = .02). This indicates that the specific group of MZ triplets that shared a chorion had a lower mean BW than all other types of triplets within the group of DZ trios.

Table 2. *Descriptive statistics for a complete group of triplets born in the Netherlands between 2000 and 2006, a subset from this group including all triplets still alive after 28 days, and the NTR sample.*

	Dutch triplets (cohort 2000 – 2006)		Dutch triplets (cohort 2000 – 2006) alive after 28 days		NTR sample with chorionicity information from PALGA	
Cohort	2000 - 2006		2000 - 2006		1987 - 2005	
Parity (% primiparous)	54.7		50.9		51.3	
	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N
BW	1745 (1681 – 1809*)	1324	1920 (1888 – 1952*)	1092	1884 (1810 – 1958*)	346
GA	32.3 (4.1)	1323	33.4 (2.8)	1092	33.6 (2.4)	346
Age mother	31.6 (4.3)	486	31.7 (4.2)	379	30.6 (3.7)	109

Note: BW = birth weight; CI = confidence interval; GA = gestational age; NTR = Netherlands Twin Register. * = for BW the 95% CI is given instead of the SD. The first two groups, the SD of triplet BW was retrieved from the perinatal registry, CI was calculated with an adjusted N size (N of families instead of N of triplets). For the third group CI was calculated directly from the data with the STATA option robust cluster.

Table 3. *Characteristics of the analyzed sample in comparison to the sample for which PALGA did not provide information on chorionicity and the mothers who could not be linked to PALGA.*

	Analyzed sample	No Chorionicity Info from PALGA	No linked to PALGA
N trios	116	60	98
Educational level mother (low; middle; high)	38.6%; 40.6%; 20.8%	49.4%; 28.2%; 22.4%	38.5%; 31.7%; 29.8%
Age mother at birth (range, Mean (SD))	21 – 41, 30.5 (3.7)	24 – 41, 31.5 (3.3)	20 – 39, 30.8 (3.6)
Population density (< 1000 persons per m^2)	52.6%	53.4%	43.9%;
Range birth cohort triples	1987 – 2005	1978 – 2006	1948 – 2006
Percentage of sample born before 1991	23.7%	17.1%	25.6%

In addition, in the last column of Table 7, within the trio, mean birth weight of the MZ pair is calculated and subtracted from the birth weight of the DZ triplet. This column shows that, within DZ trios, mean birth weight of the MZ pair was always lower when the MZ pair had shared a chorion. If not, mean birth weight of the MZ pair was about the same as the birth weight of the DZ triplet.

Discussion

This is the first study to examine the combined effect of chorionicity and zygosity on triplet birth weight. It appeared that chorionicity is the most important risk factor for low birth weight, no additional effect of zygosity on triplet birth weight was detected. This result was most clearly seen in the DZ trios. A within-family comparison in the DZ trios between a DZ individual and individuals from the MZ pair showed that sharing a chorion explains the main effect on lowering birth weight.

Based on the *triplet* zygosity and chorionicity, MZ MC triplets were found to have a lower birth weight than the MZ DC and DZ DC triplets. Due to the examination based on triplet zygosity and chorionicity rather than trio zygosity and chorionicity, the results in triplets can be compared to results in twins. Some studies in twins show the same results, while others indicate that MZ DC twins also have a lower birth weight than DZ DC twins (Lee et al., 2010; Dube et al., 2002; Loos et al., 2005). These different findings are explained by differences in study methods. First, when the effect of zygosity on birth weight is small, this effect will only be detected when sample sizes are large. In addition, to find an effect that is caused by zygosity (and not by differences in gestational age between zygosity groups), the effect of gestational age should be corrected for. Not all studies correct for factors affecting birth weight.

In this study, groups based on *trio* and on *triplet* zygosity and chorionicity were compared on gestational age, as it had been put forward that with less choria in a triplet pregnancy, gestational age is lower (Spencer et al., 2009). No evidence was found to support this effect. A lower gestational age would (partly) explain a lower birth weight in the trios of which more triplets had shared a chorion. Although no significant effects were found for lower gestational age, this characteristic was entered in the regression analysis, to correct for even small effects. Therefore, a lower gestational age could not be the explanation for the lower birth weight in the MZ MC group based on triplet chorionicity, and for the lower birth weight in the triplets from a MZ pair that had shared the chorion within the DZ trios.

In addition to gestational age, the percentages of triplets born after ART were also presented for each group based on zygosity and chorionicity. Machin and Bamforth (1996) put forward that trios born after ART have a better prognosis than spontaneously conceived triplets because in the latter group more MZ triplet pairs and trios are born. With respect to birth weight, the current study confirms that in general, triplets born after ART tend to have a higher birth weight due to the more favorable chorionicity.

Regarding birth weight discordance, it already has been established in twins that the MC are more discordant than the DC twins (Blickstein and Keith, 2004; Gielen et al., 2008; González-Quintero et al., 2003). In addition, one study showed that MZ MC and DZ DC did not differ in birth weight discordance even though both had a higher discordance than the MZ DC twins (Gielen, Derom, Derom, Vlietinck and Zeegers, 2009). Our findings in triplets suggest that,

Table 4. *Percentages of children with normal, low, very low and extremely low birth weight as a function of trio and triplet zygosity and chorionicity.*

	Trio zygosity and chorionicity						Triplet zygosity and chorionicity		
BW	MZ MC (%)	MZ DC (%)	DZ DC (%)	MZ TC (%)	DZ TC (%)	TZ TC (%)	MZ MC (%)	MZ DC (%)	DZ DC (%)
	N = 33	N = 15	N = 72	N = 5	N = 54	N = 167	N = 81	N =41	N =209
NBW: BW>2500 gram	-	-	8	20	9	6	4	10	7
LBW: 1500<BW<2500 gram	67	80	63	20	82	74	63	73	74
VLBW: 1000<BW<1500 gram	33	20	24	20	7	17	30	12	16
ELBW: BW<1000 gram	-	-	6	40	2	2	4	5	3

Note: BW = birth weight, NBW = normal birth weight, LBW = low birth weight, VLBW = very low birth weight, ELBW = extremely low birth weight.

Table 5. *Mean birth weight (gram), birth weight discordance, gestational age and ART as a function of chorionicity and zygosity.*

Chorionicity	Zygosity	N	Min. BW	Max. BW	Mean BW	95% CI	Mean BW disc*(%)	GA (SD)	%ART
MC	MZ	33	1100	2465	1749	1541 - 1956	17.4	33.1 (1.9)	0
DC	DZ	15	1060	2270	1760	1412 - 2109	22.0	33.4 (2.7)	0
DC	DZ	72	565	3090	1845	1645 - 2046	24.5	32.9 (3.1)	48
TC	TZ	5	770	2540	1546	0 - 10532	35.8	31.0 (2.8)	100
TC	TZ	54	905	2920	2069	1933 - 2205	19.7	34.3 (2.1)	16
TC	TZ	167	730	3190	1888	1781 - 1995	20.6	33.6 (2.5)	76

Note: BW = birth weight, GA = gestational age, ART = assisted reproductive techniques. * = discordance is given between heaviest and lightest triplet from a trio.

Table 6. *Triplet birth weight and gestational age as a function of triplet chorionicity and zygosity.*

Zygosity and chorionicity	N	Min. BW	Max. BW	Mean BW	95% CI	Mean BW disc*(%)	GA (SD)	% ART
MZ MC	81	757	3090	1764	1625 - 1902	16.6	33.0 (2.6)	33.8
MZ DC	41	770	2670	1982	1770 - 2194	14.3	33.9 (2.4)	15.8
DZ DC	209	565	3190	1920	1830 - 2010	19.3	33.6 (2.5)	78.8

Note: BW = birth weight, GA = gestational age, ART = assisted reproductive techniques

Table 7. *Within family comparison of DZ trios*

	Mean BW MZ pair (95% CI, N)	Mean BW DZ triplet (95% CI, N)	% diff BW DZ – MZ*
DZ DC	1774 (1573 - 1976, 48)	1987 (1763 - 2211, 24)	100
DZ TC	2043 (1897 - 2188, 36)	2121 (1913 - 2329, 18)	61.1

Note: *percentage of the trios with a heavier DZ triplet than mean birth weight of the MZ pair. DZ trios consist of 2 individuals who form a MZ pair and 1 DZ sibling. In DZ DC triplets, there were 2 choria indicating that the MZ pair shared a chorion. In DZ TC pairs all 3 individuals had their own chorion.

though the differences were not significant, birth weight discordance was lowest in the MZ DC group, some higher in the MZ MC group, and highest in the DZ DC group. These findings are in line with the hypothesis as proposed by Gielen et al. (2009).

A limitation of the current study is that PALGA did not provide chorionicity information for all the families that were linked. An important reason for this was that chorionicity information had to be extracted from the laboratory notes. Recently, PALGA has developed a standardized protocol for reporting on placenta pathology, which will be incorporated as of January 2012. As of then, new reports on placenta pathology will follow this protocol. Future collaboration with PALGA is therefore likely to result in a higher percentage of successful information gatherings.

Not all contrasts that we considered showed significant effects. In part, this might be due to sample size. This study has the largest sample so far that included both chorionicity and zygosity to examine their additional effect on birth weight in triplets. Some findings, though not significant, indicated an interesting trend, which needs to be further explored in future research. If possible, future studies should include information on chorionicity in multiple research. In twins it has been shown that low birth weight is associated with behavioral and emotional problems, and the data on this association supports a causal one (e.g. (Groen-Blokhuis, Middeldorp, Beijsterveldt and Boomsma, 2011). The effect of low birth weight (and possibly additional effects of chorionicity) therefore also needs to be considered in triplets.

In conclusion, our results suggest that chorionicity is a more important factor influencing triplet birth weight than zygosity and several mechanisms can play a role in this association. In MC twins fetal growth restriction and discordant birth weight can be caused by placental vascular anastomoses, which can result in acardiac twins, acute and chronic twin-to-twin transfusion syndrome (Nikkels, Hack and van Gemert, 2008). In addition, unequal placenta sharing can cause severe growth restriction and birth weight discordance, although only in the presence of arterio-arterial anastomoses (Hack et al., 2008). In contrast, these phenomena are not present in DC twins and discordant growth is often ascribed to differences in placental mass or differences in placental parenchymal lesions (Nikkels et al., 2008).

CHAPTER 7:

SUMMARY AND GENERAL DISCUSSION

Summary

Chapter 2 looked at Internalizing problems (INT) in adolescent twins and analyzed self-report data on anxious depression and withdrawn behavior in a bivariate genetic model. The results show that the impact of Additive genetic influences (A), Common environment shared by children from the same family (C) and unique Environment (E) on INT depend on age, thereby indicating age x genotype interaction. At age 12, the estimates of the explained variance by C for anxious depression and withdrawn behavior are significantly present, however, at age 14 and 16 genetic factors were sufficient to explain familial clustering, shared environmental effects were absent. There were no differences in the estimates of A and E between boys and girls.
Chapter 3 indicates that the teacher, school and/or class environment affects the expression of genetic factors for INT and Externalizing problems (EXT). Twins attending the same class and therefore rated by the same teacher show slightly lower levels of problem behavior and greater resemblance in problem behavior than twins who were rated by different teachers. This greater resemblance cannot be ascribed to the possibility that teachers were confusing the twins or by a bias of the teachers. Rather, the results appear to be due to a higher heritability of problem behavior assessed in children who were rated by the same teacher, which constitutes evidence for teacher/classroom x genotype interaction. In children who are exposed to the same teacher and/or classroom environment, a certain behavior can be more or be less triggered, thus teachers and the atmosphere in a classroom can truly make a difference in how children perform and behave at school.

In chapter 4, Gene-environment interaction (GxE) is investigated for INT and EXT measured at ages 3, 5 and 7 with SES (SES) and formal child care between birth and age 4. In a large sample, two competing theoretical models were tested regarding the question whether mean levels of problem behavior are influenced by formal child care and whether the influence of genetic and environmental factors on problem behavior is moderated by the attendance of formal child care. One model, the diathesis-stress model, states that higher mean levels of problem behavior are observed in risk environments, accompanied by an increase in the heritability of problem behavior in risk environments. In the other model, the bioecological model, a decrease in heritability and an increase in the mean level of problem behavior are predicted when exposed to the risk factor. Children who went to formal child care scored slightly higher on EXT, especially when these children came from a lower SES family. The impact of environmental factors was higher in the formal child care group, leading to a lower heritability. Children in the lower SES group were also more influenced by environmental effects, thus heritability was lower in those groups. The effects of child care and SES on the influence of genetic and environmental factors were most pronounced at age 7 and for EXT. With regard to the two theoretical models, these findings support the bioecological model above the stress-diathesis model, given the hypothesis that formal child care is a risk environment for problem behavior.
Chapters 5 focused on determinants of triplet birth weight because triplets are often born with a low birth weight, which is associated with disadvantageous childhood development. Gestational age, sex, zygosity and maternal smoking affect birth weight with early born, girls, monozygotic triplets, and children from smoking mothers having a lower birth weight than later born, boys, and di- or trizygotic triplets, and children of non-smoking mothers. No effects of Assisted Reproductive Technologies (ART) or maternal alcohol consumption during pregnancy were found. The resemblance in birth weight was higher in MZ triplets than in DZ triplets indicating that genetic factors are also of importance in birth weight.
Chapter 6 subsequently investigated the effect of chorionicity on birth weight of triplets. Data of 346 triplets from the NTR were successfully linked to chorionicity information in PALGA (a nationwide network and registry of histo- and cytopathology in the Netherlands). Monozygotic, monochorionic triplets had a lower mean birth weight than dizygotic, dichorionic triplets. In dizygotic trios, monozygotic pairs only had a lower mean birth weight than their dizygotic sibling triplet when the pair shared a chorion. In conclusion, sharing a chorion is a more important risk factor for low birth weight than monozygosity.

Discussion

This thesis points to the importance of analyzing multiple phenotypes, different environmental factors, and taking age into account when estimating the heritability of problem behavior. As Kendler and others (Kendler et al., 2008b; Kendler et al., 2008c; Reinke and Herman, 2002) found for INT, the influence of the genome is dynamic during childhood and adolescence. This is also true for environmental factors. The large twin registries that have followed their subjects for, for example, 25 years, such as in the Netherlands Twin Register (NTR) contribute to the disentanglement of the etiology of INT and EXT across the lifespan. This thesis had a focus that went beyond heritability and looked at genotype x environment interactions, where "environment" was broadly defined as age, teacher/classroom, formal child care and parental SES.

The findings of chapter 2 are in line with previous studies on gene x age interaction for INT problems that also showed higher heritability estimates in adolescence than in childhood and a common environmental influence only in childhood (Bergen, Gardner and Kendler, 2007; Gregory and Eley, 2007; Rapee, Schniering and Hudson, 2009; Rice, 2009). These age differences do not necessarily imply that different genes are involved, and longitudinal studies should address this question. In addition to previous research, this thesis shows that the increase in heritability occurs not during but after the onset of puberty as at age 12. Thus, when common environmental influences are still present, most of the children have already entered puberty. Furthermore, the bivariate analysis of anxious-depressive and withdrawn behavior indicated a high (0.85 to 1) genetic correlation between these phenotypes suggesting that in future genome wide association studies (GWAS), genes indicated to affect one of the two phenotypes are also likely to be involved in the other. Moreover, when including subjects in a GWAS analysis, data from individuals with measurements of only one of the two phenotypes could be pooled together to enhance the power of the analysis.

Chapter 3 and 4 investigated the effect of teacher/classroom, SES and formal child care on the influence of genes and environment on INT and EXT. Since their parents or teachers rated the twins at different ages during childhood, gene x age interaction was also taken into account. Overall, the results emphasize that both genetic and environmental effects can vary across ages and different environmental exposures and that it is not possible to draw general conclusions. This is in contrast to the conclusion of Hicks et al. (2009a and 2009b) who studied GxE interaction for several environmental risk factors in INT disorders and EXT disorders and concluded that the way genes and environment interact in a phenotype generalizes over environmental stressors. For EXT, all environmental risk factors increased the genetic variance, and, to a lesser extent, the environmental variance, which could be interpreted as, in the case of EXT phenotypes, evidence for a diathesis-stress model. On the other hand for INT, all environmental risk factors increased the environmental variance, suggesting a bioecological model. The findings in this thesis do not support a phenotype specific pattern. In the analysis of the teacher data the results show the same pattern for INT and EXT, namely a decrease in genetic variance as well as an increase in unique environmental variance when children attend a different class (and thus are exposed to different teachers). In the analysis of SES and formal child care, the environmental variance for INT and EXT was higher when children went to formal child care and when SES was lower, although the effect was more pronounced in EXT. These findings support the bioecological model, with no phenotypic specific pattern, which is thus in contrast to the Hicks et al. (2009b) paper. An important difference between the analyses in current thesis and the Hicks et al. papers was that the analyses reported here focus on 3-, 5-, 7-, 10- and 12-year-old children and the Hicks at al. paper analyzed data from adolescents at age 17. Other studies also yielded different results than the Hicks et al (2009a and 2009b) studies. Life events, for example, have also been associated with an increase of the heritability of INT, supporting the diathesis stress model (Eaves, Silberg and Erkanli, 2003; Lau and Eley, 2008; Lau et al., 2007; Silberg et al., 2001), while living in a socio-economically less advantaged neighborhood has been associated with less genetic effects on EXT, supporting a bioecological model (Legrand, Keyes, Mcgue, Iacono and Krueger, 2008; Tuvblad et al., 2006).

Detecting GxE and theorizing about the mechanisms, which lead to this phenomenon, are important not only for understanding how genes, age and environment interact and ultimately result in a phenotypic outcome. There are also more specific implications. The importance of classroom environment for internalizing and externalizing problem behavior has been discussed by (Herman, Reinke, Parkin, Traylor and Agarwal, 2009; Reinke and Herman, 2002). Chapter 3 indicates the effect of a teacher/classroom depends on a child's genetic make-up. As a consequence, the previously reported associations between, for example, positive behavior of a teacher and lower INT and EXT scores (Herman et al., 2009; Reinke and Herman, 2002), might be driven by a specific group of children and different intervention might be more suitable for other groups of children. Future studies should take child characteristics into account that can influence the effect of an intervention.

The importance of SES as a risk factor for problem behavior was already well established. A recent study also showed the enduring effect of low SES in childhood (Ramanathan, Balasubramanian and Krishnadas, 2013). Chapter 4 shows that the effect of SES on problem behavior in childhood is due to an increase in environmental variance in line with the bioecological model, which hypothesizes that risk environments mask genetic differences between children. This signifies that in the low SES category, there are more differences in individual circumstances that explain variation in problem behavior, while in the high SES category, there is less variation in the environment. More insight in the differences within the low SES category is necessary to be able to identify the children with the highest risk.

Chapter 5 shows that the same risk factors for low birth weight in singleton and twin pregnancies apply to triplet pregnancies. Since triplet pregnancies are more at risk for a complicated course, these mothers should be even more encouraged to quit smoking. Moreover, given the effect of chorionicity as shown in chapter 6, careful monitoring of monochorionic triplets is required.

Birth weight is an important determinant of childhood development and we have found that BW in triplets is 1900 grams (and gestational age is 33.7 weeks). Therefore, the next step in future research with triplets will be to establish associations with BW and the development of behavioral and emotional problems and with cognitive development. To this end, longitudinal data on behavioral and emotional problems are collected with the Child Behavior Check List (CBCL) and the Teacher Report Form (TRF), as described in the Appendix of this thesis. The NTR also collects information on CITO scores (a standardized test on Educational attainment at age 12 years) in twins and triplets (for triplets we currently have 79 scores by self-report and 37 scores via parental report). It is noteworthy that almost 50% of the surveys that are sent to parents of triplets when they have reached the age of 12 years is returned to the NTR with the remark from parents that the children have not reached 'group 8', which is the last group in elementary school in the Netherlands which nearly all children reach at the age of 12 years. When sufficient data are becoming available, we can assess in future projects the influence of BW and genetic factors on behavior in triplets and school performance.

One question to be addressed in future research involves the influence of BW on mean levels of behavioral and emotional problems and cognition. Are triplets at higher risk than singleton and twin children because they have lower GA and BW? NTR studies show that in twins there is very little difference between twins and singletons, especially when birth order within the family is taken into account, but it is not known whether this is also true in triplets (Robbers et al., 2010; de Zeeuw, van Beijsterveldt, de Geus and Boomsma, 2012). An important assumption of all studies on multiples is that the results can be generalized to singletons, who constitute the largest part of the general population, although currently about 1 in 32 children in the Netherlands are part of a multiple birth (Glasner, Beijsterveldt, Willemsen and Boomsma, 2013). Twins and triplets carry a substantial risk of preterm delivery resulting in 15–20% of all preterm births (Hall, 2003; goldenberg, Culhane, Iams and Romero, 2008). In addition, partly due to prematurely and partly due to sharing a uterus, multiples have on average a lower birth weight than singletons. In the NTR the average duration of twins pregnancies is 36.3 (2.6) weeks with an average birth weight of 2423 (542) grams in MZ twins and 36.7 (2.4) weeks with an average birth weight of 2533 (550) grams in DZ twins (Gielen et al., 2010). In singletons, prematurity and low birth weight are thought of as indicators of sub-optimal development during the gestation. Both are associated with child and adolescent INT and EXT behavior (Arnoudse-Moens, Weisglas-Kuperus, van Goudoever and Oosterlaan, 2009). In twins, their on average lower birth weight and shorter gestational age does not appear to result in an elevated risk in problem behavior in comparison with singletons. However, statistics on gestational age and birth weight in triplets are worse than in twins. Thus, the question arises how studies on triplets should be interpreted and how birth weight should be handled as possible confounding factors (i.e. do the same standards apply for triplets as they do for singletons when classifying the infants' birth weight (e.g. Low Birth Weight (LBW), Very Low Birth Weight (VLBW)). Before this question can be properly answered, factors affecting birth weight in triplets should be investigated.

A second question involves the interaction of genetic factors and BW on the development of behavioral and emotional problems and cognition in triplets. The models described in this thesis to study GxE interaction in twin data (recognizing that BW cannot be labeled as a purely environmental trait) can be extended to triplet data.

Thirdly, it may be that triplet status itself interacts with the genetic influences on behavior. For example, the rearing environment of triplets may differ more from the environment of singleton children than is the case for twins compared to singletons. (Feldman, Eidelman and Rotenberg, 2004) Feldman et al. (2004) have shown that maternal sensitivity is lower for triplets than singletons. Starting from 2012, when parents of triplets received the CBCL they were also sent an additional questionnaire which inquired about triplet well-being and the experience of parents on having and raising triplets (see appendix IV of this thesis). In addition, when triplets and twins in the NTR reach the age of 14 years, they receive surveys for the collection of self-report data (the Dutch Health and behavior questionnaire: DHBQ; Bartels et al., 2007). At that time we also approach the singleton siblings of twins and triplets. The DHBQ includes questions about the rearing environment as perceived by the twins, triplets and their singleton sibs and this, together with the longitudinal data in these groups, will constitute a valuable database for future studies, going beyond heritability studies and also including gene-environment correlation and interaction explorations.

Summary

Chapter 2 looked at Internalizing problems (INT) in adolescent twins and analyzed self-report data on anxious depression and withdrawn behavior in a bivariate genetic model. The results show that the impact of Additive genetic influences (A), Common environment shared by children from the same family (C) and unique Environment (E) on INT depend on age, thereby indicating age x genotype interaction. At age 12, the estimates of the explained variance by C for anxious depression and withdrawn behavior are significantly present, however, at age 14 and 16 genetic factors were sufficient to explain familial clustering, shared environmental effects were absent. There were no differences in the estimates of A and E between boys and girls.
Chapter 3 indicates that the teacher, school and/or class environment affects the expression of genetic factors for INT and Externalizing problems (EXT). Twins attending the same class and therefore rated by the same teacher show slightly lower levels of problem behavior and greater resemblance in problem behavior than twins who were rated by different teachers. This greater resemblance cannot be ascribed to the possibility that teachers were confusing the twins or by a bias of the teachers. Rather, the results appear to be due to a higher heritability of problem behavior assessed in children who were rated by the same teacher, which constitutes evidence for teacher/classroom x genotype interaction. In children who are exposed to the same teacher and/or classroom environment, a certain behavior can be more or be less triggered, thus teachers and the atmosphere in a classroom can truly make a difference in how children perform and behave at school.

In chapter 4, Gene-environment interaction (GxE) is investigated for INT and EXT measured at ages 3, 5 and 7 with SES (SES) and formal child care between birth and age 4. In a large sample, two competing theoretical models were tested regarding the question whether mean levels of problem behavior are influenced by formal child care and whether the influence of genetic and environmental factors on problem behavior is moderated by the attendance of formal child care. One model, the diathesis-stress model, states that higher mean levels of problem behavior are observed in risk environments, accompanied by an increase in the heritability of problem behavior in risk environments. In the other model, the bioecological model, a decrease in heritability and an increase in the mean level of problem behavior are predicted when exposed to the risk factor. Children who went to formal child care scored slightly higher on EXT, especially when these children came from a lower SES family. The impact of environmental factors was higher in the formal child care group, leading to a lower heritability. Children in the lower SES group were also more influenced by environmental effects, thus heritability was lower in those groups. The effects of child care and SES on the influence of genetic and environmental factors were most pronounced at age 7 and for EXT. With regard to the two theoretical models, these findings support the bioecological model above the stress-diathesis model, given the hypothesis that formal child care is a risk environment for problem behavior.
Chapters 5 focused on determinants of triplet birth weight because triplets are often born with a low birth weight, which is associated with disadvantageous childhood development. Gestational age, sex, zygosity and maternal smoking affect birth weight with early born, girls, monozygotic triplets, and children from smoking mothers having a lower birth weight than later born, boys, and di- or trizygotic triplets, and children of non-smoking mothers. No effects of Assisted Reproductive Technologies (ART) or maternal alcohol consumption during pregnancy were found. The resemblance in birth weight was higher in MZ triplets than in DZ triplets indicating that genetic factors are also of importance in birth weight.
Chapter 6 subsequently investigated the effect of chorionicity on birth weight of triplets. Data of 346 triplets from the NTR were successfully linked to chorionicity information in PALGA (a nationwide network and registry of histo- and cytopathology in the Netherlands). Monozygotic, monochorionic triplets had a lower mean birth weight than dizygotic, dichorionic triplets. In dizygotic trios, monozygotic pairs only had a lower mean birth weight than their dizygotic sibling triplet when the pair shared a chorion. In conclusion, sharing a chorion is a more important risk factor for low birth weight than monozygosity.

Discussion

This thesis points to the importance of analyzing multiple phenotypes, different environmental factors, and taking age into account when estimating the heritability of problem behavior. As Kendler and others (Kendler et al., 2008b; Kendler et al., 2008c; Reinke and Herman, 2002) found for INT, the influence of the genome is dynamic during childhood and adolescence. This is also true for environmental factors. The large twin registries that have followed their subjects for, for example, 25 years, such as in the Netherlands Twin Register (NTR) contribute to the disentanglement of the etiology of INT and EXT across the lifespan. This thesis had a focus that went beyond heritability and looked at genotype x environment interactions, where “environment” was broadly defined as age, teacher/classroom, formal child care and parental SES.

The findings of chapter 2 are in line with previous studies on gene x age interaction for INT problems that also showed higher heritability estimates in adolescence than in childhood and a common environmental influence only in childhood (Bergen, Gardner and Kendler, 2007; Gregory and Eley, 2007; Rapee, Schniering and Hudson, 2009; Rice, 2009). These age differences do not necessarily imply that different genes are involved, and longitudinal studies should address this question. In addition to previous research, this thesis shows that the increase in heritability occurs not during but after the onset of puberty as at age 12. Thus, when common environmental influences are still present, most of the children have already entered puberty. Furthermore, the bivariate analysis of anxious-depressive and withdrawn behavior indicated a high (0.85 to 1) genetic correlation between these phenotypes suggesting that in future genome wide association studies (GWAS), genes indicated to affect one of the two phenotypes are also likely to be involved in the other. Moreover, when including subjects in a GWAS analysis, data from individuals with measurements of only one of the two phenotypes could be pooled together to enhance the power of the analysis.

Chapter 3 and 4 investigated the effect of teacher/classroom, SES and formal child care on the influence of genes and environment on INT and EXT. Since their parents or teachers rated the twins at different ages during childhood, gene x age interaction was also taken into account. Overall, the results emphasize that both genetic and environmental effects can vary across ages and different environmental exposures and that it is not possible to draw general conclusions. This is in contrast to the conclusion of Hicks et al. (2009a and 2009b) who studied GxE interaction for several environmental risk factors in INT disorders and EXT disorders and concluded that the way genes and environment interact in a phenotype generalizes over environmental stressors. For EXT, all environmental risk factors increased the genetic variance, and, to a lesser extent, the environmental variance, which could be interpreted as, in the case of EXT phenotypes, evidence for a diathesis-stress model. On the other hand for INT, all environmental risk factors increased the environmental variance, suggesting a bioecological model. The findings in this thesis do not support a phenotype specific pattern. In the analysis of the teacher data the results show the same pattern for INT and EXT, namely a decrease in genetic variance as well as an increase in unique environmental variance when children attend a different class (and thus are exposed to different teachers). In the analysis of SES and formal child care, the environmental variance for INT and EXT was higher when children went to formal child care and when SES was lower, although the effect was more pronounced in EXT. These findings support the bioecological model, with no phenotypic specific pattern, which is thus in contrast to the Hicks et al. (2009b) paper. An important difference between the analyses in current thesis and the Hicks et al. papers was that the analyses reported here focus on 3-, 5-, 7-, 10- and 12-year-old children and the Hicks at al. paper analyzed data from adolescents at age 17. Other studies also yielded different results than the Hicks et al (2009a and 2009b) studies. Life events, for example, have also been associated with an increase of the heritability of INT, supporting the diathesis stress model (Eaves, Silberg and Erkanli, 2003; Lau and Eley, 2008; Lau et al., 2007; Silberg et al., 2001), while living in a socio-economically less advantaged neighborhood has been associated with less genetic effects on EXT, supporting a bioecological model (Legrand, Keyes, Mcgue, Iacono and Krueger, 2008; Tuvblad et al., 2006).

Detecting GxE and theorizing about the mechanisms, which lead to this phenomenon, are important not only for understanding how genes, age and environment interact and ultimately result in a phenotypic outcome. There are also more specific implications. The importance of classroom environment for internalizing and externalizing problem behavior has been discussed by (Herman, Reinke, Parkin, Traylor and Agarwal, 2009; Reinke and Herman, 2002). Chapter 3 indicates the effect of a teacher/classroom depends on a child's genetic make-up. As a consequence, the previously reported associations between, for example, positive behavior of a teacher and lower INT and EXT scores (Herman et al., 2009; Reinke and Herman, 2002), might be driven by a specific group of children and different intervention might be more suitable for other groups of children. Future studies should take child characteristics into account that can influence the effect of an intervention.

The importance of SES as a risk factor for problem behavior was already well established. A recent study also showed the enduring effect of low SES in childhood (Ramanathan, Balasubramanian and Krishnadas, 2013). Chapter 4 shows that the effect of SES on problem behavior in childhood is due to an increase in environmental variance in line with the bioecological model, which hypothesizes that risk environments mask genetic differences between children. This signifies that in the low SES category, there are more differences in individual circumstances that explain variation in problem behavior, while in the high SES category, there is less variation in the environment. More insight in the differences within the low SES category is necessary to be able to identify the children with the highest risk.

Chapter 5 shows that the same risk factors for low birth weight in singleton and twin pregnancies apply to triplet pregnancies. Since triplet pregnancies are more at risk for a complicated course, these mothers should be even more encouraged to quit smoking. Moreover, given the effect of chorionicity as shown in chapter 6, careful monitoring of monochorionic triplets is required.

Birth weight is an important determinant of childhood development and we have found that BW in triplets is 1900 grams (and gestational age is 33.7 weeks). Therefore, the next step in future research with triplets will be to establish associations with BW and the development of behavioral and emotional problems and with cognitive development. To this end, longitudinal data on behavioral and emotional problems are collected with the Child Behavior Check List (CBCL) and the Teacher Report Form (TRF), as described in the Appendix of this thesis. The NTR also collects information on CITO scores (a standardized test on Educational attainment at age 12 years) in twins and triplets (for triplets we currently have 79 scores by self-report and 37 scores via parental report). It is noteworthy that almost 50% of the surveys that are sent to parents of triplets when they have reached the age of 12 years is returned to the NTR with the remark from parents that the children have not reached 'group 8', which is the last group in elementary school in the Netherlands which nearly all children reach at the age of 12 years. When sufficient data are becoming available, we can assess in future projects the influence of BW and genetic factors on behavior in triplets and school performance.

One question to be addressed in future research involves the influence of BW on mean levels of behavioral and emotional problems and cognition. Are triplets at higher risk than singleton and twin children because they have lower GA and BW? NTR studies show that in twins there is very little difference between twins and singletons, especially when birth order within the family is taken into account, but it is not known whether this is also true in triplets (Robbers et al., 2010; de Zeeuw, van Beijsterveldt, de Geus and Boomsma, 2012). An important assumption of all studies on multiples is that the results can be generalized to singletons, who constitute the largest part of the general population, although currently about 1 in 32 children in the Netherlands are part of a multiple birth (Glasner, Beijsterveldt, Willemsen and Boomsma, 2013). Twins and triplets carry a substantial risk of preterm delivery resulting in 15–20% of all preterm births (Hall, 2003; goldenberg, Culhane, Iams and Romero, 2008). In addition, partly due to prematurely and partly due to sharing a uterus, multiples have on average a lower birth weight than singletons. In the NTR the average duration of twins pregnancies is 36.3 (2.6) weeks with an average birth weight of 2423 (542) grams in MZ twins and 36.7 (2.4) weeks with an average birth weight of 2533 (550) grams in DZ twins (Gielen et al., 2010). In singletons, prematurity and low birth weight are thought of as indicators of sub-optimal development during the gestation. Both are associated with child and adolescent INT and EXT behavior (Arnoudse-Moens, Weisglas-Kuperus, van Goudoever and Oosterlaan, 2009). In twins, their on average lower birth weight and shorter gestational age does not appear to result in an elevated risk in problem behavior in comparison with singletons. However, statistics on gestational age and birth weight in triplets are worse than in twins. Thus, the question arises how studies on triplets should be interpreted and how birth weight should be handled as possible confounding factors (i.e. do the same standards apply for triplets as they do for singletons when classifying the infants' birth weight (e.g. Low Birth Weight (LBW), Very Low Birth Weight (VLBW)). Before this question can be properly answered, factors affecting birth weight in triplets should be investigated.

A second question involves the interaction of genetic factors and BW on the development of behavioral and emotional problems and cognition in triplets. The models described in this thesis to study GxE interaction in twin data (recognizing that BW cannot be labeled as a purely environmental trait) can be extended to triplet data.

Thirdly, it may be that triplet status itself interacts with the genetic influences on behavior. For example, the rearing environment of triplets may differ more from the environment of singleton children than is the case for twins compared to singletons. (Feldman, Eidelman and Rotenberg, 2004) Feldman et al. (2004) have shown that maternal sensitivity is lower for triplets than singletons. Starting from 2012, when parents of triplets received the CBCL they were also sent an additional questionnaire which inquired about triplet well-being and the experience of parents on having and raising triplets (see appendix IV of this thesis). In addition, when triplets and twins in the NTR reach the age of 14 years, they receive surveys for the collection of self-report data (the Dutch Health and behavior questionnaire: DHBQ; Bartels et al., 2007). At that time we also approach the singleton siblings of twins and triplets. The DHBQ includes questions about the rearing environment as perceived by the twins, triplets and their singleton sibs and this, together with the longitudinal data in these groups, will constitute a valuable database for future studies, going beyond heritability studies and also including gene-environment correlation and interaction explorations.

RISICOFACTOREN VOOR PROBLEEMGEDRAG TIJDENS DE JEUGD EN ADOLESCENTIE: STUDIES IN TWEELINGEN EN DRIELINGEN

Dit proefschrift bundelt vijf studies die risicofactoren voor probleemgedrag bij kinderen en jongeren onderzoeken. Er wordt onderscheid gemaakt tussen *Internaliserende problemen*, zoals angstig, depressief en teruggetrokken gedrag, en *externaliserende problemen*, zoals normafwijkend en agressief gedrag. Uit eerdere gedragsgenetische studies was al bekend dat tijdens de kindertijd en adolescentie, probleemgedrag erfelijk is, waarbij de mate van erfelijkheid afhankelijk is van de leeftijd en het geslacht van het kind. Het was minder goed bekend hoe omgevingsfactoren de erfelijkheid van probleemgedrag beïnvloeden of hoe het effect van omgevingsfactoren kan afhangen van genen. Vaak gedragen kinderen zich in de ene situatie anders dan in de andere. Het is mogelijk dat dit komt doordat in een risico-omgeving het genetisch profiel van het kind meer tot uiting komt. Wanneer een risico-omgeving zorgt voor gemiddeld meer probleemgedrag plus een *hogere* erfelijkheid van dit gedrag, dan wordt dit in de gedragsgenetica beschreven als het diathesis-stress model. Dus de aanleg voor ziekte (diathesis) komt tot uiting bij stress. Het tweede model is het bio-ecologische model. In dit model wordt gesteld dat in een risico-omgeving meer probleemgedrag, maar juist een *lagere* erfelijkheid van dit gedrag gevonden wordt. De eerste drie studies in dit proefschrift onderzoeken deze gen-omgevingsinteractie, waarbij omgeving breed gespecificeerd is als leeftijd, de schoolse omgeving (leerkracht/klas), formele kinderopvang, en sociaaleconomische status (SES) van de ouders.

In de laatste twee studies in dit proefschrift worden factoren die van invloed zijn op drielinggeboortegewicht onderzocht. De meeste kinderen komen als eenling ter wereld en bij deze kinderen is er een verband aangetoond tussen laaggeboortegewicht, prematuriteit, en probleemgedrag. De geboorte van een drieling is veel unieker. De meeste drielingen worden prematuur geboren en hebben een laag tot zeer laag geboortegewicht. Wat zijn factoren die het geboortegewicht bij drielingen in gunstige of ongunstige zin kunnen beïnvloeden? Er is gekeken naar de effecten van roken en alcoholconsumptie tijdens de drielingzwangerschap, vruchtbaarheidsbehandelingen, zwangerschapsduur, sekse, geboortevolgorde, zygositeit en de chorioniciteit van de drielingen.

In al deze studies is gebruik gemaakt van gegevens verzameld door het Nederlands Tweelingen Register (NTR). Het NTR volgt al meer dan 25 jaar meerlingen met het doel een bijdrage te leveren aan het onderzoek naar de effecten van genen en omgeving op probleemgedrag tijdens de jeugd en adolescentie en aan het onderzoek naar de groei en ontwikkeling van meerlingen.

Hoofdstuk 2. In het tweede hoofdstuk worden zelfrapportage data van adolescente tweelingen over angstig/depressief en teruggetrokken gedrag geanalyseerd met een bivariaat genetisch model. Uit de resultaten blijkt dat de erfelijkheid van het gedrag leeftijdsafhankelijk is, waarbij de erfelijkheid toeneemt tussen 12 en 16 jaar. Dit is een aanwijzing voor gen-leeftijdsinteractie. Daarnaast bleek dat de omgeving die tweelingen delen (zoals bijvoorbeeld de familie of de sociaaleconomische status van het gezin) nog wel van invloed is op internaliserend gedrag op leeftijd 12, maar niet meer op latere leeftijd. Er was geen verschil in de erfelijkheid van angstig/depressief en teruggetrokken gedrag tussen jongens en meisjes. De scores op angstig/depressief gedrag en teruggetrokken gedrag nemen bij zowel jongens als meisjes toe. Wel scoren meisjes gemiddeld hoger dan jongens.

Hoofdstuk 3. De studie beschreven in het derde hoofdstuk laat zien dat de leerkracht en/of de klas waarin een kind zit tijdens de basisschool, van invloed is op de mate waarin het genetisch profiel van een kind tot expressie komt. Tweelingen die in dezelfde klas zitten en door dezelfde leerkracht werden beoordeeld vertoonden meer overeenkomst in probleemgedrag dan kinderen die door verschillende leerkrachten werden beoordeeld. De grotere overeenkomst kon niet verklaard worden doordat leerkrachten de kinderen door elkaar haalden of door een mogelijk vooroordeel van de leerkracht. Gen-omgevingsinteractie was wel een passende verklaring. Voor kinderen in dezelfde klas was de erfelijkheid van probleemgedrag rond de 70 procent, voor kinderen in verschillende klassen was dit rond de 30 procent.

Hoofdstuk 4. In dit hoofdstuk is onderzocht of formele kinderopvang en sociaaleconomische status van de ouders een effect hebben op de erfelijkheid van internaliserend en externaliserend gedrag op de leeftijden 3, 5 en 7 jaar. Verder is er gekeken of de erfelijkheid toe- of afnam in de risico-omgeving, dat wil zeggen, was er sprake van een diathesis-stress dan wel bio-ecologisch model. Het bleek dat kinderen die naar formele kinderopvang gingen iets hoger scoorden op externaliserend gedrag. Het effect van sociaaleconomische status was echter iets sterker, kinderen uit families met een laag sociaaleconomische status scoorden hoger op externaliserend gedrag.

Verder werd er gevonden dat kinderen uit families met een laag sociaaleconomische status meer beïnvloed werden door omgevingsinvloeden, dus de erfelijkheid was lager in die groepen. Deze effecten waren het sterkst op leeftijd 7. De resultaten ondersteunen een bio-ecologisch model.

Hoofdstuk 5. In de studie beschreven in dit hoofdstuk worden risicofactoren voor een laag geboortegewicht van drielingen onderzocht. Er werden effecten gevonden van roken tijdens de zwangerschap, de zwangerschapsduur, sekse, genetische aanleg en de zygositeit van het kind. Kinderen van rokende moeders, die vroeg geboren, monozygoot en van vrouwelijk geslacht zijn hebben het laagste geboortegewicht. Er zijn geen effecten gevonden van vruchtbaarheidsbehandelingen en alcoholconsumptie tijdens de zwangerschap op het geboortegewicht van drielingen. Monozygote drielingen leken in geboortegewicht meer op elkaar dan dizygote drielingen, wat een indicatie is voor een effect van erfelijke factoren op het geboortegewicht van drielingen.

Hoofdstuk 6. In het zesde hoofdstuk is vervolgens het effect van chorioniciteit op drielinggeboortegewicht onderzocht. Data van 346 NTR drielingen zijn gekoppeld aan chorioniciteitsdata van stichting PALGA (een Nederlands nationaal netwerk en register voor histo- en cytopathologie). Monozygote en monochorionische drielingen hadden een lager gemiddeld geboortegewicht dan dizygote en dichorionische drielingen. Bij de dizygote trio's hadden de kinderen die een monozygoot paar vormden gemiddeld een lager geboortegewicht dan hun dizygote broertje of zusje. Geconcludeerd kan worden dat het delen van een chorion een belangrijkere risicofactor is voor laag geboortegewicht bij drielingen dan monozygositeit.

Discussie

Uit dit proefschrift blijkt dat wanneer erfelijkheid als risicofactor voor probleemgedrag tijdens de jeugd en adolescentie wordt onderzocht, het belangrijk is om onderscheid te maken tussen het soort probleemgedrag (internaliserend dan wel externaliserend), effecten van de omgeving, en de leeftijd van de kinderen. De invloed van het genoom op probleemgedrag is dynamisch tijdens de jeugd en adolescentie. Dit geldt ook voor de omgeving. Tweelingenregisters zoals het NTR die meerlingen al meer dan 25 jaar volgen dragen bij aan het onderzoek naar hoe genen en omgeving van invloed zijn op groei en ontwikkeling tijdens de jeugd. De focus van dit proefschrift ging verder dan de simpele erfelijkheidsbepaling en keek naar gen-omgevingsinteractie, waarbij de “omgeving” breed gespecificeerd was als leeftijd, leerkracht /klas, formele kinderopvang en sociaaleconomische status van de ouders.

De bevinding uit hoofdstuk 2 dat de erfelijkheid van internaliserend gedrag hoger is in de adolescentie dan op 12-jarige leeftijd en dat op die leeftijd de invloed van de omgeving die tweelingen delen nog wel aanwezig is, komt overeen met eerdere studies. Een nieuwe bevinding is dat de toename van de erfelijkheid *na* 12-jarige leeftijd en de start van de pubertijd plaatsvindt. Dus, gedeelde omgeving is ook na de start van de pubertijd nog steeds van invloed. Daarnaast is in dit proefschrift aangetoond dat genen die van invloed zijn op angstig/depressief gedrag grotendeels overlappen met genen die teruggetrokken gedrag beïnvloeden.

In hoofdstuk 3 en 4 worden gen-leeftijdsinteractie en gen-omgevingsinteractie onderzocht. De mate waarin genetische effecten van invloed zijn op probleemgedrag tijdens de jeugd is afhankelijk van leeftijd en omgeving. Daarom is het dus niet verstandig om over leeftijden en verschillende omgevingen te generaliseren. In eerdere studies zijn voor externaliserend gedrag aanwijzingen gevonden voor een diathesis-stress model. Daarnaast rapporteerden andere studies voor internaliserend gedrag al indicaties voor een bio-ecologisch model. In dit proefschrift werd gevonden dat de invloed van genen relatief kleiner is bij kinderen die naar formele kinderopvang zijn geweest en uit gezinnen komen met een lagere sociaaleconomische status. Dit werd gevonden voor zowel internaliserend als externaliserend gedrag, maar het effect was het sterkst voor externaliserend gedrag. Deze bevinding ondersteunt een bio-ecologisch model en is ook in overeenstemming met eerdere bevindingen met betrekking tot de sociaaleconomische status van de buurt waarin kinderen wonen.

Onderzoek naar gen-omgevingsinteractie is niet alleen van belang voor het begrip van de *oorzaak* van probleemgedrag tijdens de jeugd. Hoofdstuk 3 toont aan dat een praktische implicatie is dat het effect van een leerkracht of klas op probleemgedrag afhankelijk is van een kinds genetisch profiel. Met deze kennis lijkt het nuttig om op basisscholen interventies tegen probleemgedrag te ontwikkelen die vormgegeven worden afhankelijk van het kind, en er inzicht is dat één uniforme benadering mogelijk niet zal werken.

Het was al bekend dat een lage sociaaleconomische status van het gezin waarin kinderen opgroeien een risicofactor is voor meer probleemgedrag. Hoofdstuk 4 toont aan dat effecten van sociaaleconomische status deels lopen via een toename van omgevingsinvloeden op probleemgedrag en niet via het meer tot uiting komen van een genetische kwetsbaarheid. Voor kinderen uit families met een laag sociaaleconomische status is er meer variatie in omgevingsfactoren die probleemgedrag beïnvloeden dan bij kinderen uit families met een hoog sociaaleconomische status. Vervolgonderzoek moet zich richten op de identificatie van deze specifieke risicofactoren binnen de groepen met een lagere sociaaleconomische status.

Hoofdstuk 5 toont aan dat risicofactoren voor laag geboortegewicht van één- en tweelingen ook van toepassing zijn voor laag geboortegewicht van drielingen. Omdat een drielingzwangerschap risicovol is, zouden rokende moeders nog sterker moeten worden aangeraden te stoppen met roken. Daarnaast werd in hoofdstuk 6 gevonden dat monochorioniciteit een voorspeller is van lager geboortegewicht, dus zou het verstandig zijn om de monochorionische drielingen nog beter onder controle te houden.

In eenlingen is aangetoond dat een laag geboortegewicht een predictor is voor problemen tijdens de ontwikkeling. In dit proefschrift is beschreven dat het geboortegewicht van drielingen gemiddeld 1900 gram is. Het is daarom van belang om in toekomstig onderzoek te bepalen hoe het geboortegewicht van drielingen en de ontwikkeling van probleemgedrag en cognitie geassocieerd zijn. Om dit te onderzoeken is het noodzakelijk om longitudinaal het gedrag van drielingen te volgen. De appendix van dit proefschrift beschrijft de start van de dataverzameling bij drielingen die meedoen binnen het NTR. Aan hun ouders is gevraagd de “Child Behavior Check List” in te vullen. Leerkrachten hebben de “Teacher Report Form” ontvangen. Daarnaast heeft het NTR ook de CITO score voor drielingen op leeftijd 12 verzameld. Het is interessant dat hierbij duidelijk is geworden dat op leeftijd 12, 50% van de ouders van drielingen rapporteert dat de kinderen nog niet in groep 8 zitten. Wanneer genoeg data beschikbaar zijn kan onderzocht worden hoe het geboortegewicht en mogelijk genetische factoren van invloed zijn op het gedrag en de schoolprestaties van drielingen.

Andere belangrijke vragen voor toekomstig onderzoek is of het lage geboortegewicht van drielingen van invloed is op het gemiddeld niveau van probleemgedrag en cognitie van deze kinderen. Zijn drielingen een risicogroep door hun laaggeboortegewicht en prematuriteit? Een belangrijke aanname in het tweelingenonderzoek is dat resultaten gebaseerd op tweelingen ook naar de algehele (voornamelijk éénling) populatie kan worden gegeneraliseerd. Tweelingen en drielingen hebben een verhoogd risico op prematuriteit. Vijftien tot twintig procent van alle prematuur geboren kinderen zijn een meerling. Daarnaast hebben meerlingen door zowel de prematuriteit als door het delen van de baarmoeder gemiddeld een lager geboortegewicht. Een tweelingzwangerschap duurt gemiddeld 36 weken met een afgerond gemiddeld geboortegewicht van 2400 gram voor monozygote en 2500 gram voor dizygote tweelingen. Voor eenlingen geldt dat een laag geboortegewicht een indicator is voor een minder goed verlopen ontwikkeling tijdens de zwangerschap. Zowel korte zwangerschapsduur als een laag geboortegewicht zijn geassocieerd met later probleemgedrag. Voor tweelingen lijkt het lagere geboortegewicht en de kortere zwangerschapsduur geen nadelige effecten met zich mee te brengen. Maar drielingen zijn nog lichter en worden nog vroeger geboren dan tweelingen. Het is dus de vraag hoe het geboortegewicht mogelijk een vertekenend beeld geeft wanneer probleemgedrag van drielingen wordt onderzocht.

Een tweede vraag is hoe genetische factoren en geboortegewicht, probleemgedrag en de cognitie van drielingen beïnvloeden. De modellen die gen-omgevingsinteractie onderzoeken en die beschreven zijn in dit proefschrift zouden ook op data van drielingen kunnen worden toegepast, hoewel geboortegewicht natuurlijk niet als een pure omgevingsfactor kan worden beschouwd.

Een derde vraag is of het ook mogelijk is dat drielingstatus interacteert met genetische invloeden op probleemgedrag en cognitie. De thuissituatie en/of het gedrag van een ouder naar een drieling zou bijvoorbeeld anders kunnen zijn dan voor een eenling. Vanaf 2012 hebben ouders van drielingen, naast de Child Behavior Check List, ook een vragenlijst over de opvoeding, ouderschap en welbevinden van de drieling ontvangen (zie appendix IV van dit proefschrift). Daarnaast ontvangen drielingen vanaf leeftijd 14 zelf een vragenlijst. Vanaf die leeftijd worden broertjes en zusjes van twee- en drielingen ook benaderd voor deelname aan het vragenlijstonderzoek. Gevraagd wordt onder andere hoe meerlingen en hun broertjes en zusjes hun leefomgeving ervaren. Deze data zullen een belangrijke bijdrage leveren aan toekomstig onderzoek dat verder gaat dan slechts bepalingen van de erfelijkheid van probleemgedrag tijdens de jeugd en adolescentie, maar ook gen-omgevingscorrelatie en interactie onderzoekt