Behavior problems, cognition, and hormones

A longitudinal-genetic study in childhood

Meike Bartels

This research was financially supported by The Netherlands Organization for Scientific Research (575-5-012). The Netherlands Organization for Scientific Research (NWO: R 56-467) and the Stichting Simonsfonds (SF053-iz) provided travel grants to facilitate collaboration with dr. E.J.C.G. van den Oord, at the Virginia Institute of Psychiatric and Behavioral Genetics, Department of Psychiatry, Medical College of Virginia/ Virginia Commonwealth University, Richmond, Virginia, USA.

ISBN 90-9016490-1
Printed by [OPTIMA] grafische communicatie, Rotterdam
Cover Eric van Rossum
Lay-out Meike Bartels

VRIJE UNIVERSITEIT

Behavior problems, cognition, and hormones

A longitudinal-genetic study in childhood

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. T. Sminia,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de faculteit der Psychologie en Pedagogiek
op donderdag 9 januari 2003 om 15.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Meike Bartels

geboren te Gouda

promotor : prof.dr. D.I. Boomsma
copromotor : dr. E.J.C.G. van den Oord

Leescommissie : prof.dr. J.K. Buitelaar
prof.dr. P.T. Cohen-Kettenis
prof.dr. H.A. Delemarre-van de Waal
prof.dr. C. Kirschbaum
prof.dr. J.F. Orlebeke
prof.dr. J. Terwel

TABLE OF CONTENTS

1

General introduction

Information on the origin of individual differences in development during childhood may be considered as valuable to gain insight into the etiology of human variation. In studying the development of children two major phenotypes are often considered: cognitive abilities and childhood psychopathology. For both childhood psychopathology and cognitive abilities the importance of genetic and environmental influences at various ages is well established. Further, it is widely shown that both phenotypes are stable over the years. Less information is available on the genetic and environmental influences on the development of cognitive abilities and childhood psychopathology over the years. For a comprehensive picture of developmental processes underlying child psychopathology and cognition the mechanisms that explain continuity and change need to be understood. Genetic and environmental influences, for instance, can yield similar or distinct influences on the developmental process.

One of the problems in previous studies on stability and change in development during childhood is the phenotypic character of the data used. Stability and change of a trait throughout development, based on data collected in samples of unrelated subjects, are mostly expressed as correlations over time. Most previous studies have relied on these phenotypic correlations, so they could not distinguish whether continuity was caused by genetic or environmental influences or both. A more fundamental problem of analyzing phenotypic data is that genetic and environmental factors may display different developmental patterns. A mixture of different developmental patterns is not distinguishable at phenotypic level, so that using phenotypic data only could lead to false conclusions.

The key to these fundamental problems is studying genetic related individuals. One of the most important methods to study the etiology of human variation is the classical twin design (See Chapter 2). In short, monozygotic (MZ) twins derive from a single zygote and are therefore genetically identical or nearly identical (Petronis, 2001; Martin *et al.*, 1997). A possible way to explain differences between two members of a MZ twin pair, as indicated by less than perfect MZ twin correlations ($r_{MZ} < 1$), are environmental effects that are not shared. Nonshared environmental influences (E) denote the impact of all environmental factors influencing only one of the subjects being studied, such as an illness, diseases, trauma, experiences at school or relationships with peers. Dizygotic (DZ) twins develop from two distinct zygotes and share on average 50% of their segregating genes, like 'ordinary' brothers and sisters. Differences between two members of a DZ twin pair can result from nonshared environmental influences as well as genetic differences. A higher observed resemblance of MZ versus DZ twin pairs ($r_{MZ}>r_{DZ}$) is an indication for genetic influences (A) on the trait under investigation (Martin *et al.*, 1997). The twin design also allows the study of environmental influences that are shared by members of a twin pair.

Shared environmental factors (C) will create differences between families and make family members relatively more similar. Possible examples are socioeconomic level, religion, or style of parenting. Complex traits such as the ones studied in psychology are likely to be influenced by multiple genes each with a small effect. It can be shown that if these genes act in an additive manner, the DZ twin correlation is half the MZ twin correlation (Falconer and Mackay, 1996). Shared environmental influences are implied for traits where $r_{DZ} > \frac{1}{2} r_{MZ}$. The study of genetically related individuals, though, provides more than estimates of genetic and environmental influences on complex traits. The statistical properties of these data yield unique information on fundamental questions in psychology such as rater (dis)agreement and developmental mechanisms not available in phenotypic designs, in which traits are measured in genetically unrelated subjects.

Possible moderators for individual differences in cognition and problem behavior are hormones such as cortisol. These possible moderating effects are partly based on a phenomenon dubbed 'fetal programming' (for a review see: Welberg and Seckl, 2001; Matthews, 2000). In the prenatal period the development of the brain is influenced by hormones, secreted by the pituitary and the gonads (Collaer and Hines, 1995; Sikich and Todd, 1988). These early effects are referred to as 'programming effects' because of their possible influence on the development and structure of the brain. In other words, individual differences in pre- or postnatal cortisol levels could be an indirect cause of individual differences in psychopathology or cognition (see Chapter 11).

The main objective of this study is to investigate the developmental patterns of cognitive abilities and problem behavior. To this end distinct longitudinal patterns are considered and the cross sectional and longitudinal influences of multiple raters, for assessment of problem behavior, on the estimates of genetic and environmental influences in childhood psychopathology are investigated. Finally, since cortisol levels are considered as a possible moderator for individual differences in cognition and psychopathology, an overview of previous studies on individual differences is cortisol levels is given and the individual differences in cortisol levels in 12-year old children are investigated.

Childhood psychopathology

Prevalence

Behavioral/emotional problems are common among children. In a Dutch sample of 2227 children, aged 4 to 18 years, the prevalence of behavioral problems is about 40% (Verhulst *et al.*, 1996). This number illustrates that problem behaviors in children present a public health problem that cannot be ignored. Further, Internalizing problems (anxious/depressed behavior, withdrawn behavior) are more prevalent in girls than boys,

especially in adolescence, and Externalizing problems (aggressive behavior, rule breaking behavior) are more prevalent in boys than girls.

In developmental psychopathology there is considerable interest in the study of how problem behaviors develop over time. So far studies on the development of childhood psychopathology mainly focus on phenotypic correlations over time and persistence of problem behavior to predict stability and change (Verhulst and Van der Ende, 1992a; 1992b; Ghodsian *et al.*, 1980; Richman *et al.*, 1982; Graham and Rutter, 1973). These studies indicate a substantial degree of continuity in problem behaviors. In two prevalence studies in representative samples of Dutch children, the 8-year stability of Child Behavior Checklist (CBCL) total problem scores in children, initially aged 4 to 10 years, was r=0.48 (Verhulst *et al.*, 1985a; 1985b; 1989; 1997; 1999; Verhulst and Akkerhuis, 1986; Verhulst and Van der Ende, 1995). This means that across the rather long period of eighth years, nearly 25% of the variance in CBCL total problem scores could be explained by the initial CBCL total problem scores. From a categorical perspective, it was found that nearly 40% of those who could be regarded as deviant, could still be regarded as deviant eighth years later. Stability of Internalizing problems was nearly equally strong as the stability of Externalizing problems.

Genetics

Only a few studies have used data from genetically related individuals to disentangle the genetic and environmental influences on continuity and change in the development of problem behaviors or problem behavior related disorders. Van der Valk and colleagues (2002, submitted) used a two-wave behavior genetic model to estimate genetic, shared environmental and nonshared environmental contributions to stability and change of Internalizing and Externalizing Problems at ages 3 and 7 years in a sample of Dutch twins. For Externalizing problems the estimated influences of additive genetic, shared and nonshared environmental factors remained relatively constant over the years. The phenotypic stability (r=.54) was explained for 55% by genetic factors. Shared environmental influences were mostly stable, while nonshared environmental influences were mostly age specific. For Internalizing problems additive genetic influences decreased while nonshared environmental influences increased over the years. The phenotypic stability (r=.38) was for 66% explained by genetic factors. In a developmental study in sibs, half sibs and cousins, by Van den Oord and Rowe (1997), the continuity of problem behaviors was entirely explained by genetic and shared environmental factors. Nonshared environmental factors only showed age specific effects, influencing changes in children's problem behaviors.

One of the major drawbacks in the previous phenotypic studies is the lack of ability to sort out the etiology of stability and change in problem behavior throughout development. These previous phenotypic studies could not distinguish between genetic and environmental influences on the developmental process. Previous studies in genetically related individuals made a distinction between genetic and environmental factors causing stability and change. These different sources of variances, though, may display a distinct developmental pattern, for instance a factor or simplex structure (explained in Chapter 2). An important feature of the present longitudinal twin study is the possibility to investigate the developmental pattern of genetic and environmental sources of variance independently. Unique in this study is the large longitudinal data set with information on childhood psychopathology at age 3, 7, 10, and 12.

Childhood psychopathology and the Netherlands Twin Register

The Netherlands Twin Register (NTR), kept by the Department of Biological Psychology at the Vrije Universiteit in Amsterdam, was established in 1987 (Boomsma *et al.*, 1992; Boomsma, 1998, Boomsma *et al.*, 2002). Young twins and multiples are recruited a few weeks or months after their birth. Currently around 50% of all newborn multiples in The Netherlands are registered. The young twins are recruited with the help of a commercial organization (Felicitas B.V.) that visits parents of newborns at home. After the parents send in the signed registration card, they receive the first questionnaire. When this questionnaire is returned, the family is registered with the NTR and will be send a questionnaire every two years, and a newsletter yearly. Over the years, there is an increase in the number of registrations (defined by the return of a completed first questionnaire), which corresponds with the increase in the number of twin and multiple births in The Netherlands. Between 1989 and 2000, the relative number of multiple births in The Netherlands increased from 13.2‰ (a total of 2250 twin and multiple births in 1987) to 17.8‰ (a total of 3639 twin and multiple births in 2000; data CBS). Since the questionnaire collection is a continuous process, Table 1.1 lists the questionnaires that have been returned by parents at each age till the first of June 2002.

The main focus of the questionnaires is health and behavior problems. For the study of development of Internalizing and Externalizing problem behavior questionnaire data on psychopathology collected at ages 3, 7, 10, and 12 in birth cohorts 1986 till 1993 are used, as indicated by the shaded cells in Table 1.1. Further, the questionnaires used at these ages contain a serie of items on parental beliefs about zygosity, twin resemblance for physical characteristics and confusion of twins by parents, family members and strangers. In combination with information on zygosity based on blood group and DNA

polymorphisms, these data have been used to assign zygosity to same-sex twins (Rietveld *et al.*, 2000).

Table 1.1.

Number of twin pairs for whom questionnaires are returned by parents

Cohort	0	2	3	5	7	10	12 years
1986	155	109	82	101	89	103	103
1987	940	810	634	592	534	580	544
1988	993	815	673	635	601	586	522
1989	1022	795	800	710	720	636	504
1990	1189	945	945	861	802	745	157
1991	1235	988	900	867	837	713	
1992	1336	1072	973	872	906	210	
1993	1477	1161	1080	956	937		
1994	1533	1190	1000	928	703		
1995	1588	1132	1057	1017	57		
1996	1749	1261	1151	963			
1997	1664	1196	1156	118			
1998	1779	1480	475				
1999	1685	1282	4				
2000	1426	118					
2001	916						
Total	20687	14354	10930	8620	6186	3573	1830

Note: Two questionnaires are sent to parents at ages 3, 5, 7, 10, and 12 years. The numbers in the table indicate that at least one of these is returned. The shaded cells indicate the cohorts and age groups used in this study.

Data on psychopathology are assessed by a worldwide-accepted questionnaire, the *Child Behavior Checklist* (CBCL; Achenbach, 1991; 1992). Mother and father ratings were collected by making use of age appropriate questionnaires. The CBCL 2/3 is used for parents to score the behavioral and emotional problems of their 3-year-old children. It consists of 100 items that are scored by the parents on a 3-point scale based on the occurrence of the behavior during the preceding 2 months: 0 if the problem item was not true, 1 if the item was somewhat or sometimes true, and 2 if it was very true or often true. Dutch syndrome scales and comparability with the syndrome scales as developed by Achenbach (1992) are reported by Koot and colleagues (1997). In the present project the two broadband scales Internalizing and Externalizing are analyzed (Chapter 3 to 6). In the

CBCL 2/3 the Internalizing scale consists of the Anxious and Withdrawn/Depressed subscales. The Externalizing scale consists of the Aggressive, Oppositional, and Overactive subscales.

For twins at age 7, 10, and 12 the CBCL 4-18 was used for parents to score the behavioral and emotional problems of their children. It consists of 120 problem items that are scored by the parents on, the above mentioned, 3-point scale based on the occurrence of the behavior during the preceding 6 months. The syndrome scales were composed according to the 1991 profile (Achenbach, 1991). Dutch syndrome scales and comparability with the syndrome scales as developed by Achenbach are reported in Verhulst *et al.* (1996). In this manual the two broadband scales Internalizing and Externalizing are analyzed. In the CBCL 4-18, the Internalizing scale consists of the Anxious/Depressed, Somatic Complaints and Withdrawn subscales. The Externalizing scale consists of the Aggressive and Rule Breaking Behavior subscales.

Multiple informants

The used approach to quantify children's problem behavior by asking the parents to score behavioral and emotional problems on behavioral questionnaires has its advantages and its disadvantages. Parents have the advantage that they observe their children over long periods and can witness both frequent and rare behaviors. Additionally, using more than one rating will give more reliable results by decreasing measurement error. Disadvantages are the disagreement between mother and father ratings and the costs and efforts to send questionnaires to multiple persons.

Traditionally, data from mothers have been used since information from fathers has been difficult to obtain, and once obtained, difficult to interpret, given the often low levels of agreement between mother and father reports. A meta-analysis by Achenbach, McConaughy, and Howell (1987) showed a mean correlation of .60 between maternal and paternal ratings of the same child. This underscores that parents are able to assess their child's behavior, for if parental ratings would reflect nothing but error the correlation between their ratings would probably be low. However, a high correlation does not necessarily imply that parents are assessing the same phenotype in their children (Hewitt *et al.*, 1992). The correlations may be high even when the parents are assessing different behaviors in their children, because the parental correlation may predominate over the variance specific to a given parent. Conversely, different forms of rater bias and unreliability may lower the correlation between parents even though parents may be assessing exactly the same phenotype in their children. Generally, disagreement can have elements of accuracy or bias. For example, accuracy may arise if one parent is privy to a child's behavior in ways unavailable to other informants, e.g., as when a child confides with

his or her mother about personal problems. In this case, the mother, but not the father, would have unique access to data on the son or daughter's feelings. Bias could arise if the parents' own traits influenced ratings (a projection bias), or if parents exhibited response biases (e.g. stereotyping, employing different normative standards, or having certain response styles, i.e. judging problem behaviors more or less severely).

From a clinical point of view, it remains a struggle to determine what to do with the disagreement. Is it best to assume that there is 'one best informant', that one parent is 'more reliable than the other', or that parents present a unique viewpoint on his or her child, thus providing unique and valuable information to be used in assessment? Accordingly, it is difficult to draw firm conclusions about the processes underlying the (dis)agreement between parental ratings on the basis of correlations alone.

To study agreement and disagreement between parental ratings two models are considered (see Chapter 2). The so-called Rater Bias and Psychometric models combine data of mothers and fathers and can be used to estimate genetic and environmental influences taking agreement and disagreement between parents into account (Neale and Stevenson, 1989; Hewitt *et al.*, 1992) . The Rater Bias model assumes that parents assess exactly the same behaviors in the child (common behavioral view) and that they share a common understanding of the behavioral descriptions. This may apply when both parents are equally confronted with the behaviors shown by the child (for instance at home). Disagreement between the raters is regarded as error, resulting from rater bias and/or unreliability. Sources of rater bias are stereotyping, employing different normative standards, or having certain response styles, i.e. judging problem behaviors more or less severely. Because these types of bias may differ between raters, they may also lead to disagreement between raters. Unreliability can become an important source of disagreement when raters cannot give an accurate description about relevant behaviors. For instance, it has been suggested that parents may be relatively insensitive to affective disturbances in children (Angold *et al.*, 1987).

In addition to assessing similar aspects of the child's behavior, the Psychometric model assumes that each parent assesses a unique aspect of the child's behavior. This will occur when the parent also observes the child in distinct situations where they are exposed to distinct samples of the behavior (unique behavioral view). For instance, the parent who usually brings the child to school may also be more familiar with the child's behavior outside the home. Moreover, each parent may interact differently with the child (Achenbach *et al.*, 1987). These unique interactions between a parent and a child may allow each parent to provide additional information about the child's behavior, apart from the information on which they both agree. Disagreement in this model does not merely arise due to unreliability and/or rater bias, but also because each parent contributes, from his

own perspective, different but valid information on the child's functioning. The Psychometric model tests this by examining whether there are significant child genetic effects on the parents' unique behavioral views. If the behaviors uniquely rated by the parents are shown to be influenced by the genotype of the child, the parent must have been assessing 'real' unique behavioral views. For error and/or unreliability cannot cause the systematic effects necessary for the model to estimate genetic influences.

Only a few studies have collected data on psychopathology rated by both parents and therefore were able to employ models that incorporated rater differences. Rowe and Kandel (1997) collected the CBCL completed by mothers and fathers for their oldest two offspring (aged 9 to 17) in 76 families. They did not fit either Psychometric or Rater Bias models. Still, their results, based on an 'individual view-shared view' model, showed that the parental ratings contained a substantial shared behavioral view. Simonoff *et al.* (1995), in a study of 282 twin pairs aged 8 to 16, also found evidence in favor of a shared behavioral view for antisocial behaviors. However, from their analyses they could not determine what underlay the shared parental view and described it as due to a shared set of expectations of the parents against which both twins were rated. Hewitt *et al.* (1992) applied both the Rater Bias and Psychometric model on parental ratings of the Internalizing scale (CBCL) for 983 twin pairs. They found that both for their prepubertal cohort (8 to 11 years) and for their pubertal cohort (12 to 16 years) the Psychometric model fitted the data better than the Rater Bias model. Hewitt and colleagues concluded that for the Internalizing scale, mothers and fathers rate the same phenotype in their children (i.e. have a shared behavioral view). Unique genetic influences were found, implying that the rater differences reflected the existence of real unique behavioral views and not just error and bias. Further insight into issues of rater bias is presented by van der Valk and colleagues (2001; 2002). Rater bias models and Psychometric models were tested on a large group of 3-and 7-year-old Dutch twins. As in the previous studies, the Psychometric model fitted the data significantly better at both ages. Thus these studies indicated that disagreement between parental ratings is partly caused by mothers and fathers assessing different aspects of the child's behavior.

Contrast effects

In studying the etiology of childhood psychopathology using twin pairs and parental ratings the effects of contrast effects need to be discussed. In general, contrast effects can be considered as a social interaction between siblings (Carey, 1986; Eaves, 1976) or an effect introduced by the rater (Neale and Stevenson, 1989). In the former case, the behavior of one twin has a certain effect on the behavior of his or her co-twin. This effect can be either cooperative or competitive. In the latter case, when parents are asked to

evaluate and report upon the problem behavior of their children, they may very well compare the twins' behavior against one another, despite instructions on the questionnaire form. In this way, one twin becomes some kind of standard by which the behavior of the co-twin is rated. A significant contrast effect is implied when MZ variances and DZ variances are heterogeneous. In addition to heterogeneity of MZ and DZ variances, the presence of a contrast effect leads to a pattern of MZ and DZ correlations that is inconsistent with additive genetic sources of variances. They decrease both MZ and DZ correlations, but DZ correlations to a greater extent, thus mimicking interaction between alleles at the same locus (dominance). The implication of a rater contrast effect in diagnosis and research is clear. Bias can lead to misdiagnoses in the clinical setting and the inclusion of false positives or exclusion of true cases from gene searching efforts, both of which are undesirable. If a gene finding study is designed to select discordant twin pairs and concordant twin pairs, the former group would be over-represented and the latter group would be under-represented due to maternal rater bias.

Cognitive abilities

Heritability of cognition has been studied extensively, both in adults and in children. Many behavior genetic studies yield the largely consistent result that genetic differences account for at least 50% of the observed variability in cognition in adults (e.g. Bouchard and McGue, 1981; McCartney *et al.*, 1990; Bratko, 1996; Rijsdijk *et al.*, 1997, 1998; Alarcón *et al*, 1998, 1999, Posthuma *et al.*, 2000). It is also well established that the genetic influences on cognitive functioning increase throughout development, whereas influences of common environment decrease (e.g. Skodak and Skeels, 1949; Wilson, 1983; Labuda *et al.*, 1986; Fulker *et al.*, 1988; Loehlin *et al.*, 1989; McCartney *et al.*, 1990; McGue *et al.*, 1993; Boomsma, 1993; Plomin *et al.*, 1997; Boomsma and Van Baal, 1998; Alarcón, 1998, 1999).

Far less is known about the developmental genetics of cognitive abilities. A few longitudinal studies have focused on the influences of genes and environment on cognitive development rather than cognition at specific ages. New genetic influences at different ages and a common factor for shared environmental influences have been found (Colorado Adoption Project; e.g. Plomin and DeFries, 1985; Louisville Twin Study; e.g. Wilson, 1983; Eaves *et al.*, 1986).

In addition to participating in the longitudinal questionnaire studies, a smaller sample of twins was invited to take part in an experimental study on the development of cognitive abilities. The development of cognitive abilities is investigated using Full-scale IQ scores at age 5, 7, 10, and 12. The use of four measurement occasions enables us to distinguish between a simplex structure and a common factor structure as the underlying pattern for

genetic and environmental influences on the development of cognition. Further the association between IQ at all ages and a test of scholastic achievement (CITO) is explored.

Psychometric IQ

At age 5, 7, and 10 the children were tested with the Revised Amsterdamse Kinder Intelligentie Test (RAKIT) (Bleichrodt *et al*, 1984). Six subtests, with age-appropriate items, were employed to assess cognitive functioning. Raw subtest total scores are corrected for age and transformed into standardized scores with a mean of 15 and a standard deviation of 5. The total IQ score is based on the combination of these transformed subtests with a mean of 100 and standard deviation of 15. The standardization is based on a population sample of Dutch 6 to 11-year-old children and the same standardization is used for boys and girls. For further details on this well-known Dutch intelligence test see Rietveld *et al.* (2000). At age 12 the twins completed the full version of the WISC-R, Dutch version (Van Haasen *et al.*, 1986). The WISC-R consists of 12 subtests, 6 mainly verbal and 6 mainly non-verbal. The subtest scores are standardized, with a mean of 100 and a standard deviation of 15, based on results of same-aged children in the Netherlands and the same standardization is used for boys and girls. Addition of the twelve standardized subtest scores results in Full-scale IQ.

Scholastic achievement (CITO)

In 2000, the NTR started collecting the results of a national test of educational achievement (CITO) from all registered 12-year old twins. 85% of all Dutch schools yearly administer this test in the final class of elementary school. The main purpose of this test is to select for different levels of high school education. A standardized CITO score was collected for 1495 children, who took the CITO in 1998, 1999, 2000 or 2001.

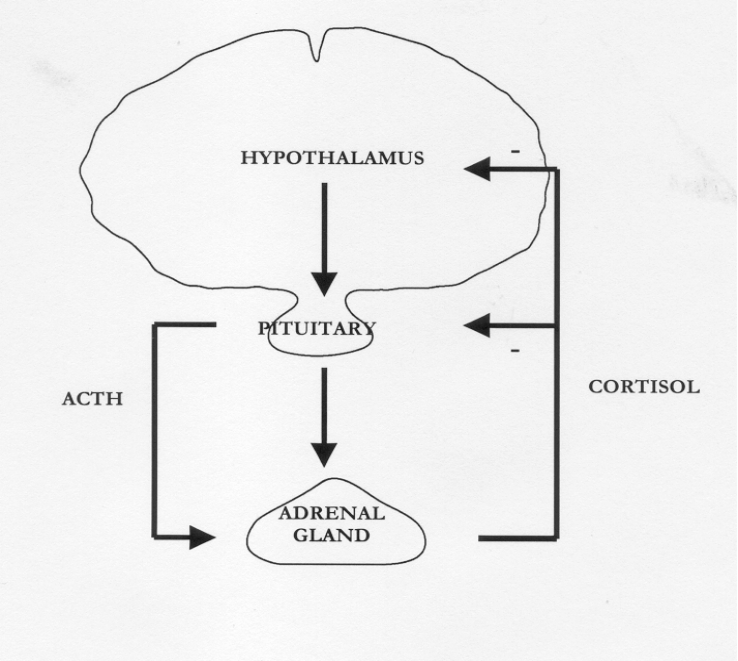
The CITO consists of 240 multiple-choice items assessing four different intellectual skills: Language, Mathematics, Information Processing, and World Orientation. Each performance scale contains 5 or 6 subscales, with a total of 60 multiple-choice questions. In 2001 the test slightly changed with respect to the distribution of the questions resulting in 60 questions for Mathematics and World Orientation, 90 questions for Language and 30 questions for Information Processing. Together the performance scales result in a standardized score between 501 and 550. The test is administered on three consecutive days in January or February when the children are in the final class of elementary school. The CITO data were collected by mail from the teacher, after informed consent from the parents or by mail from the parents as a question in a questionnaire on the child's behavior at age 12.

Cortisol

Cortisol, a glucocorticoid, is a steroid hormone secreted by the outer cortex of the adrenal gland. Its secretion is stimulated by ACTH (adrenocorticotrophic hormone), produced in the pituitary in response to corticotrophin-releasing hormone (CRH), a product from the neurons in the paraventricular nucleus of the hypothalamus (Figure 1.1.).

Figure 1.1.

A schematic representation of the HPA-axis.



In the characteristic diurnal rhythm of plasma cortisol level typically 10-15 well-defined pulses of variable amplitude are observed, with a morning maximum, declining levels throughout the daytime, a period of low concentrations generally centered around midnight, and an abrupt rise after the first few hours of sleep (Weitzman, 1971). Further, plasma cortisol release is tightly regulated through negative feedback at the pituitary, hypothalamus and hippocampus (Kovacs *et al.*, 1987; Jacobson and Sapolsky, 1991). Strength of this feedback signal strongly varies with time of day (Dorin *et al.*, 1996; Huizinga *et al.*, 1998; Young *et al.*, 1998), contributing to the characteristic diurnal rhythm in plasma cortisol levels. After its release, the major proportion of cortisol binds to the plasma proteins corticosteroid binding globulin (CBG or transcortin) and albumin, which

prevent the hormone from penetrating the membranes of their target cells. Only, about 3-5% of the total cortisol is the unbound, biologically active fraction. Salivary cortisol measurements always reflect the biological active free form. Salivary free cortisol is approximately 70% of that of serum free cortisol because of conversion of cortisol to cortisone in the salivary glands. However, there is a strong relationship between cortisol levels extracted from saliva and from blood (Riad-Fahmy *et al*; 1982; Kirschbaum and Hellhammer 1994; Aardal and Holm 1995).

Only a few attempts have been made to estimate the impact of genetic and environmental factors on the regulation of cortisol levels. These studies point to the direction of moderate genetic contributions to different aspects of cortisol measures. Thus, Maxwell et al. (1969) and Meikle et al. (1988) reported evidence of moderate genetic effects on basal cortisol levels in females and males, respectively. More recently, Inglis et al. (1999) has reported a heritability of 46% in morning plasma cortisol samples. Diurnal cortisol profiles have also been shown to be genetically affected (Linkowski et al., 1993). Furthermore, significant heritabilities have been found in the cortisol stress-response and in the cortisol response to awakening (Kirschbaum et al., 1992; Wüst et al., 2000). To date, each of these studies was done in adults and plenty of remarks can be put on the reliability of these studies (for a review see Bartels et al., 2002; Chapter 9).

For the determination of cortisol levels in the smaller sample that participated in the longitudinal IQ design, saliva was collected in 1999/2000 when the twins were 12 years old. To this end Salivettes were sent to the participants by mail and the twin pairs collected their saliva at home, following a written instruction. The samples were collected at prescribed times and, importantly, at the same time for both children of a twin pair. On the first day the first sample (S11) was taken in the morning just before getting up (still lying in bed) (mean time 0728H), the second (S21) sample was taken at least half an hour after getting up but before going to school (mean time 0817H), the third sample (S31) was taken before lunch (mean time 1234H), and the fourth sample (S41) was taken in the evening (mean time 2032H). On the second day the same schedule was adapted for four repeated samples (S12, S22, S32, S42)(mean times 0735H, 0826H, 1232H, 2034H).

Outline of this thesis

Chapter one, the introductory chapter, gives an overview of the traits under investigation and results of previous studies are presented in short. In chapter two the background methodology of the study of individual differences is discussed and a solid biometrical basis for the classical twin design will be provided. Further, the longitudinal mechanisms and the multiple rater models considered throughout this thesis are discussed in detail.

Chapter three through six will focus on childhood psychopathology. In chapter three the influences of genetic and environmental influences on the developmental pattern of Internalizing and Externalizing problem behavior will be disentangled. In chapter four and five the use of multiple raters (parents) in assessing problem behavior in children is considered. Agreement and disagreement between parents is investigated for parental ratings of their 10 and 12-year-old children. Comparison of the results of this study to the results of comparable studies in 3-and 7-year-old Dutch twins gives the opportunity to disentangle 'real' behavioral development from changes in rater effects. To gain insight into in the developmental pattern of agreement and disagreement in parental ratings a longitudinal psychometric is developed and described in chapter six.

Chapters seven and eighth will focus on cognitive abilities. Chapter seven describes the longitudinal mechanism in the development of cognitive abilities from age 5 to 12 years. The influences of genetic and environmental factors on stability and change are estimated. The influences of genetic and environmental influences on the association between cognition at age 5, 7, 10, and 12 and a test of scholastic achievement (CITO) are disentangled in chapter eighth. It is investigated whether cognitive ability at age 5 is a predictor of scholastic achievement at age 12.

Since cortisol is considered as a moderator to explain individual differences in cognition and psychopathology, in chapter nine studies on the heritability of cortisol are reviewed and the methods of the different studies are critically evaluated. A power analysis is conducted to estimate the number of twin pairs required to reliably estimate the genetic and environmental influences on basal cortisol levels. Chapter ten describes a study on the individual differences in salivary cortisol in twelve-year-old children. The possible background hypothesis of the association between cortisol, psychopathology and cognition is discussed in chapter eleven and preliminary results on the correlation between cortisol, psychopathology and cognition are presented.

In the general discussion (chapter twelve), the empirical results for the chapters three to ten are discussed and integrated with existing literature. Finally, future directions in the study on individual differences in development during childhood and adolescence are considered.

2

Individual differences

Behavior genetics comprises the study of individual differences. Quantitative genetic theory states that every individual's phenotype is made up of genetic and environmental contributions. No phenotype in psychology will be entirely determined by genetic effects, so we should always expect an environmental effect, which also includes measurement error, on the phenotype P. In general the phenotypic value (P) can be expressed as a function of the genotypic value (G) and environmental deviation (E):

$$P = G + E \qquad (eq.\ 2.1)$$

Almost everything that can be measured or counted in humans shows variation around the mean value for the population. The aim of behavioral genetic studies is to disentangle the sources of this variation. Continuous observed variation might be attributed to genetic variation and environmental variation. In short, environmental variation results in variation in the phenotype when different aspects of the environment have differential effects on that phenotype. Genotypic variation causes phenotypic variation when different alleles of one or more genes differently affect the phenotype. The variance of a phenotype (V_P) is given by:

$$V_P = g^2 V_G + e^2 V_E \qquad (eq.2.2)$$

The coefficients g and e are population parameters that represent the strength of the relation between the measured phenotype and the latent (unmeasured) factors G and E. This equation does not include the term G×E, and thereby assumes no interaction between the genetic and the environmental effects (see section on G×E and GE-correlation). G and E are latent factors, in other words, the genotypes and environments are not measured directly. There are two ways to define the individual contribution of the latent factors to the total variance. First, the factor loadings can be fixed to 1, so:

$$\text{if } g=e=1$$
$$V_P = V_G + V_E \qquad (eq.\ 2.3)$$

in which V_G is the genetic variance and V_E is the environmental variance.

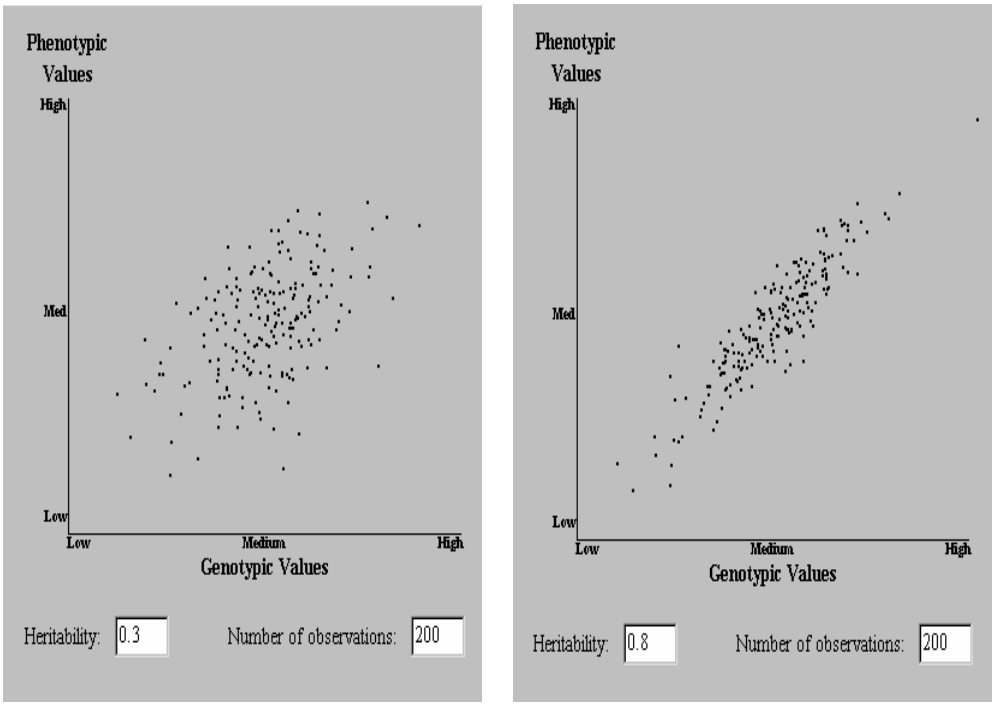
Second, when the latent factors are defined to have unit variance, a squared factor loading represents the variance explained by that specific factor, so:

$$\text{if } V_G=V_E=1$$
$$V_P= g^2 + e^2 \qquad (eq.\ 2.4)$$

in which g^2 represents the genetic variance and e^2 represents the environmental variance.

Figure 2.1.

The scatter plots demonstrate the relationship between genotypic values and phenotypic values for different heritabilities (G. Carey, 2002).



The proportion of the phenotypic variance attributable to genetic variance is known as the *heritability* (h^2). Because heritability is a proportion, its numerical value will range from 0.0 (genes do not contribute at all to phenotypic individual differences) to 1.0 (genes are the only reason for individual differences). For human behavior, almost all estimates of heritability are in the moderate range of .30 to .60, however for some phenotypes estimates as high as .80 are found (e.g. Posthuma *et al.*, 2001). The scatter plots in Figure 2.1 give some indication of the extent to which genetic individual differences contribute to individual differences in actual, observed behavior. It can be seen that if the heritability is high (e.g. 0.8; right part of Figure 2.1), there is a strong relationship between genotypic values and phenotypic values. In other words, a large part of the phenotypic individual differences is accounted for by genotypic individual differences. If the heritability is low (e.g. 0.3; left part of Figure 2.1) the relationship between genotypic values and phenotypic

values is low. In other words environmental individual differences rather than genetic individual differences will account for a large part of the phenotypic individual differences.

Genetic variation

Genetic information is present in nearly every cell of an organism and resides on long strands of deoxyribonucleic acid (DNA) called chromosomes. Humans have two copies of a chromosome, one of maternal origin and one of paternal origin. In total humans have 23 pairs of chromosomes. DNA sequences that encode for particular products (proteins and RNAs) are referred to as genes, and their chromosomal locations are called loci. Different variants of a gene are called alleles and each individual has two copies. If the two alleles of an individual are identical, the individual is said to be homozygous. If the two alleles are not identical, the individual is said to be heterozygous.

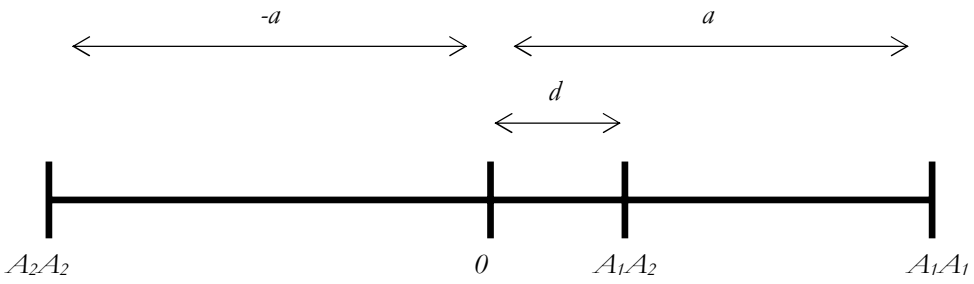
Gregor Mendel began the science of genetics with his landmark studies concerning the inheritance of the physical traits of the garden pea *Pisum sativum*, in which he studied for instance round versus wrinkled seeds (Mendel, 1866). Mendel formulated two principles of heredity that have become known as "Mendel's laws". Mendel's laws pertain to the behavior of genes and alleles during sexual reproduction. Mendel's first law, the Principle of Segregation, states that during formation of gametes the two alleles of a gene segregate separate from each other such that half of the gametes carry one allele and half carry the other allele. Mendel's second law, the Principle of Independent Assortment, states that during formation of gametes the segregation of alleles of one gene occurs independently of the segregation of alleles of other genes. A monogenetic model, though, cannot explain genetic influences on most kinds of behavior. Fisher (1918) extended Mendel's single/ two locus system to a multi-locus system. He indicated that each gene in the polygenetic model segregates according to Mendelian rules, and that small effects of these polygenes can lead to a continuous range of measured traits by summing the genetic effects of all contributing loci, and possibly also including interaction between loci (Fisher, 1918; Philips, 1998).

Within a population some genes appear in one form only, so causing no variation among individuals. For other genes, though, many different alleles may exist. For simplicity suppose a gene, A, with two different alleles A_1 and A_2. With these two alleles the genotypes may be symbolized A_1A_1, A_1A_2, and A_2A_2. The frequency of each of these genotypes (*the genotype frequencies*) is determined by the proportion of individuals that belong to each genotype. By convention, allele A_1 has a frequency p, while allele A_2 has a frequency q, so that $p + q = 1$. If the population is in Hardy-Weinberg equilibrium, the genotype frequencies are p^2, pq, and q^2 for A_1A_1, A_1A_2, and A_2A_2, respectively. The genotypic effect (*genotypic value*), that defines the measurable effect, of genotype A_1A_1 is

called "*a* ", the effect of genotype A_1A_2 is "*d* ", and the effect of genotype A_2A_2 is "*-a* ". The scaling of the three genotypes is shown is Figure 2.2.

Figure 2.2.

Scale of genotypic values (picture taken from Falconer and Mackay, 1996)



The origin, or point of zero value, on this scale is the midway between the values of two homozygotes. The value *d* of the heterozygote depends on the degree of dominance. If there is no dominance, d=0; if A_1 is dominant over A_2, *d* is positive, and if A_2 is dominant over A_1, *d* is negative. If allele A_1 is completely dominant over allele A_2, effect *d* equals effect *a*.

The genotypes, frequencies, and genotypic values are shown in Table 2.1, and from these we can calculate the mean. The genotypic contribution to the population mean of a trait is the sum of the products of the frequencies and the genotypic values of the three different genotypes (Falconer and Mackay, 1996).

Table 2.1

The genotypes, frequencies, and genotypic values for a gene A, with two different alleles A_1 and A_2.

Genotype	Frequency (f_i)	Genotypic value	Frequency × Value
A_1A_1	p^2	a	p^2a
A_1A_2	$2pq$	d	$2pqd$
A_2A_2	q^2	$-a$	$-q^2a$

The population mean is:

$$M = p^2a + 2pqd - q^2a$$
$$= a(p - q) + 2pqd \qquad (eq.\ 2.5)$$

The contribution of this locus to the population means consist of a component attributable to the homozygotes (A_1A_1 or A_2A_2), a ($p - q$), and a component attributable to the heterozygotes (A_1A_2), $2pqd$. If there is no dominance (d=0), the second term is zero, and the mean is a direct function of the gene frequency; M= a ($p - q$)= a(*1-2q*). If there is complete dominance ($d = a$), the mean is proportional to the square of the gene frequency; M=$a(p-q)+2pqa = a\,(1-2q^2)$.

However, most complex traits, like psychopathology or cognitive abilities, are not influenced by single loci. The genetic effects are assumed to be a result of the combination of the effects at several loci. With additive combination, then, the population mean resulting from the joint effects of several loci is the sum of the contributions of each of the separate loci (Falconer and Mackay, 1996), thus:

$$\mu = \sum a\,(p - q) + 2\sum pqd \qquad (eq.\ 2.6)$$

In studying individual differences we are, besides the mean of a population, interested in the variance. It reflects the dispersion of values around a mean. In studying individual differences it is important to define the genotypic effect of a certain genotype in terms of a deviation from the population mean. For the genotypes A_1A_1, A_1A_2, and A_2A_2, the genotypic effect were previously assigned a, d, and $-a$ respectively. The genotypic effects in terms of a deviation from the population mean (see equation *2.5*) are thus:

$$\begin{aligned} A_1A_1 &= a - (a\,(p - q) + 2pqd) \\ &= 2q\,(a - pd) \end{aligned} \qquad (eq.\ 2.7)$$

$$\begin{aligned} A_1A_2 &= d - (a\,(p - q) + 2pqd) \\ &= a\,(q - p) + d\,(1 - 2pq) \end{aligned} \qquad (eq.\ 2.8)$$

$$\begin{aligned} A_2A_2 &= -a - (a\,(p - q) + 2pqd) \\ &= -2p\,(a + qd) \end{aligned} \qquad (eq.\ 2.9)$$

The total genetic variance (V_G) can be computed based on the standard formula for the variance:

$$\sigma^2 = \sum f_i\,(x_i - \mu)^2 \qquad (eq.\ 2.10)$$

where f_i denotes the frequency of the genotype i (see Table 2.1), x_i denotes the mean of that genotype and μ denotes the population mean as specified in equation *2.5*. In the simple two-allele system the total genetic variance is:

$$\begin{aligned} V_G &= p^2 [\, 2q\,(a - pd)\,]^2 + 2pq [\; a\,(q - p) + \\ &\quad d\,(1 - 2pq)\,]^2 + q^2 [\, -2p\,(a + qd)\,]^2 \\ &= 2pq\,[a+d(q-p)]^2 + (2pqd)^2 \end{aligned} \qquad (eq.\ 2.11)$$

The genotypic variance (V_G) can be further divided in additive genetic variance (A) and non-additive genetic variance. Additive genetic variance (A) represent the sum of the effects of alleles of all loci that influence the trait. Non additive effects concern interactions between alleles, which can occur in two ways. Dominant genetic effects (D) stem from the summation of the interaction between two alleles at the same locus. Epistatic variance reflects the interaction between alleles at different loci. In human populations the effects of epistasis are difficult to estimate in the absence of control over breeding or environmental condition, and will henceforth be ignored. The total genetic variance is thus:

$$V_G = V_a + V_D \qquad (eq.\ 2.12)$$

In which:

$$V_A = 2pq\,[a+d(q-p)]^2 \qquad (eq.\ 2.13)$$

and

$$V_D = (2pqd)^2 \qquad (eq.\ 2.14)$$

Environmental variation

Environmental variance by definition embraces all variation of non-genetic origin. It can have a great variety of causes. If no real measures of environment are collected then it is possible to make some very important statements about the structure of the environment, by using the biometrical genetical approach, which relies only on the complex pattern of twin resemblance. Environmental variance can be decomposed in environmental variance shared by members of a family (V_C) and nonshared environmental variance (V_E), variance that is unique to a certain individual. The total environmental variance is thus:

$$V_{ENVIRONMENT} = V_C + V_E \qquad (eq.\ 2.15)$$

Any environmental factor that creates differences between families and make family members relatively more similar can be considered as shared environmental influences, for instance socioeconomic level, religion, or style of parenting. In studies of twins reared together, the shared environment is expected to contribute to the correlation of both MZ and DZ twins. Nonshared environmental influences are idiosyncratic experiences that contribute to differences between members of the same family, such as an illness, diseases, trauma, or unique relationships with peers. Measurement error is also captured in the nonshared environmental influences.

In some studies influences of shared environment are found. The finding of these significant influences, however, could be the results of nonrandom mating (Assortative mating). Assortative mating refers to nonrandom mating that results in similarity between spouses. In human populations, the first indication of assortative mating is often a correlation between the phenotypes of mates. Besides the so-called passive elements of mate selection (e.g. type and length of education, social class, area of residence), active personal preferences for physical and psychological attributes, including IQ, may play a role in mate selection. These mating elements will induce positive assortative mating. Assortative mating is important for genetic research for two reasons. First, assortative mating increases genetic variance in a population. In other words, positive assortative mating increases variance in that the offspring differ more from the average than they would if mating were random. Even if spouse correlations are modest, assortative mating can greatly increase genetic variability in a population, if its effects accumulate generation after generation. Assortative mating is also important because it affects estimates of heritability. Positive assortative mating increases the resemblance between fraternal or dizygotic twins because it renders the parents of these twins more similar than they would be if there was no assortment. Identical or monozygotic twins, however, are already at the point of maximum genetic resemblance, and are thus unaffected by positive assortative mating (Plomin *et al.*, 2000). As a result the genetic effects of assortative mating will artificially inflate estimates of the shared environmental influences. This means, in turn, that estimates of the genetic component based primarily on the difference between MZ correlations and DZ correlations will tend to be biased downwards in the presence of assortative mating. The resolution of the mechanisms of assortment is beyond the capabilities of the classical twin design, although multivariate studies that include the spouses of twins, or their parents may be capable of resolving the complex issues (see, e.g. Heath *et al.*, 1985).

Genotype-Environment effects

Several factors defy the simple separation of genetic and environmental effects, for instance, genotype-environment correlation and genotype-environment interaction. These can normally be neglected without seriously affecting the conclusions drawn from partitioning the variance, but it is important to acknowledge their existence and to know what the consequences of neglecting them are.

Genotype-environment correlation (rGE) refers to genetic effects on individual differences in liability to exposure to particular environmental circumstances, so it reflects a nonrandom distribution of environments among different genotypes. In other words, what seem to be environmental effects can reflect genetic influences because these experiences are influenced by genetic differences among individuals. Equation *2.2* is true only if environmental deviations and genotypic values are uncorrelated. When a correlation is present the phenotypic variance is increased by twice the covariance of genotypic values and environmental deviations and equation *2.2* becomes:

$$V_P = V_G + V_E + 2cov_{GE} \qquad (\textit{eq. 2.16})$$

Because of the fact that the covariance is in practice unknown, it is best regarded as part of the genetic variance. This is because the non-random aspects of the environment are a consequence of the genotypic value and so an individual's environment can be thought of as part of its genotype (Falconer and Mackay, 1996). rGE adds to the phenotypic variance for a trait, but is difficult to detect the overall extent to which phenotypic variance is due to the correlation between genetic and environmental effects (Plomin *et al.*, 1977). Christopher Jencks (1972) spoke of the 'double advantage' phenomenon in the context of cognitive ability and education. Individuals who begin life with the advantage of genes which increase their ability relative to the average may also be born into homes that provide them with more enriched environments, for instance being more committed to learning and teaching. This double advantage model can even be expanded to a triple advantage model because of the possibility that children, who are born with genes, which increase their ability and live in an enriched environment, are also more capable of profiting from this environment.

Genotype-environment interaction (G×E) refers to the genetic control of sensitivity or susceptibility to differences in the environment. In other words different genotypes respond differently to the same environment (Eaves, 1984; Mather and Jinks, 1977; Falconer and Mackay, 1996; Boomsma and Martin, 2002). The contribution of G×E to the overall population variance is typically smaller than the main effects of G and E even in controlled experiments using extreme environments. An obvious example of G×E is

that of inherited disease resistance. Genetically susceptible individuals will be free of disease as long as the environment does not contain the pathogen. Resistant individuals will be free of the disease even in a pathogenic environment. One approach for detecting G×E interaction is to estimate components of phenotypic variance conditional on environmental exposure (Eaves, 1982). Because some genotypes may be more sensitive to environmental differences than others, to some extent the environmental variance is a property of the genotype. But the source of the variation is environmental and not genetic. The interaction between genetic effects and nonshared environment, G×E, will contribute to the total variance but not to the resemblance of twin pairs. In other words, this interaction term will be confounded with nonshared environmental effects (Eaves *et al.*, 1977). If, however, G×E was interaction between genes and shared environmental influences, thus G×C , assuming its absence will result in overestimation of the effect of genes on the phenotype, as well as in overestimation of the influences of the shared environment on the phenotype. The separate detection of these two biased effects in the presence of genes by shared environmental interaction necessitates the inclusion of twins reared apart (Eaves *et al.*, 1977; Heath *et al.*, 2002).

The Classical Twin Design

The total phenotypic variance (V_p) of a trait is, in absence of genotype-environment correlation and genotype-environment interaction, the sum of the additive genetic, dominance genetic, shared environmental, and nonshared environmental variance and can be denoted with the equation:

$$V_P = V_A + V_D + V_C + V_E \qquad (eq.\ 2.17)$$

where V_P = phenotypic variance
V_A = additive genetic variance
V_D = dominant genetic variance
V_C = shared environmental variance
V_E = nonshared environmental variance

or in line with equation *2.4*:

$$V_P = a^2 + d^2 + c^2 + e^2 \qquad (eq.\ 2.18)$$

Note that in the classical twin design the effects of dominant genetic influences and shared environmental influences are confounded. Dominant genetic effects will increase

differences between MZ and DZ covariances, whereas shared environmental variance will decrease the differences. So from this point onwards all genetic effects are considered to be additive. The decomposition of the total variance will be:

$$V_P = a^2 + c^2 + e^2 \qquad (eq.\ 2.19)$$

A powerful tool to disentangle these genetic and environmental influences is to study genetically related individuals. Family studies might give a first impression of familial aggregation, but they can not distinguish between genetic and shared environmental effects. Similarities between family members may be created either by genetic relatedness or by sharing the same family environment. A method that solves this problem, is the classical twin design. Monozygotic (MZ) twins derive from a single zygote and therefore two individuals of a MZ twin pair are genetically identical. Dizygotic (DZ) twins develop from two distinct zygotes and share on average 50% of their genes, like 'ordinary' brothers and sisters. Hence, the only possible way to explain variation between two members of a MZ twin pair are environmental effects that are not shared by those two. So the covariance between two members of a MZ twin pair can be formulated as:

$$COV_{MZ} = a^2 + c^2 \qquad (eq.\ 2.20)$$

Conversely, the variation between two members of a DZ twin pair could result from different genes and/or nonshared environmental influences. DZ twins share on average half of their alleles, i.e. ½ a^2. The expectation for the covariance in DZ twin is:

$$COV_{DZ} = ½\ a^2 + c^2 \qquad (eq.\ 2.21)$$

Accordingly, the difference in relatedness between MZ and DZ twin pairs gives information about the strength of the genetic and environmental influences on the trait under investigation (Martin *et al.*, 1997). It further allows the separation of environmental influences into those of the environment shared by members of a family and those unique for each individual, the nonshared environmental influences.

Univariate design

The expectations for the resemblance between MZ and DZ twins (eq. *2.20* and eq. *2.21*) can be summarized in a path diagram (Figure 2.3). When we scale the (latent) factors by fixing their variance to one, these individual contributions can be obtained by taking the

quadratic forms of the path coefficients. By rules of path analysis the phenotypic variance (V_{T1} and V_{T2}) of twin 1 and twin 2 are:

$$V_{T1} = V_{T2} = a^2 + c^2 + e^2 \qquad (eq.\ 2.22)$$

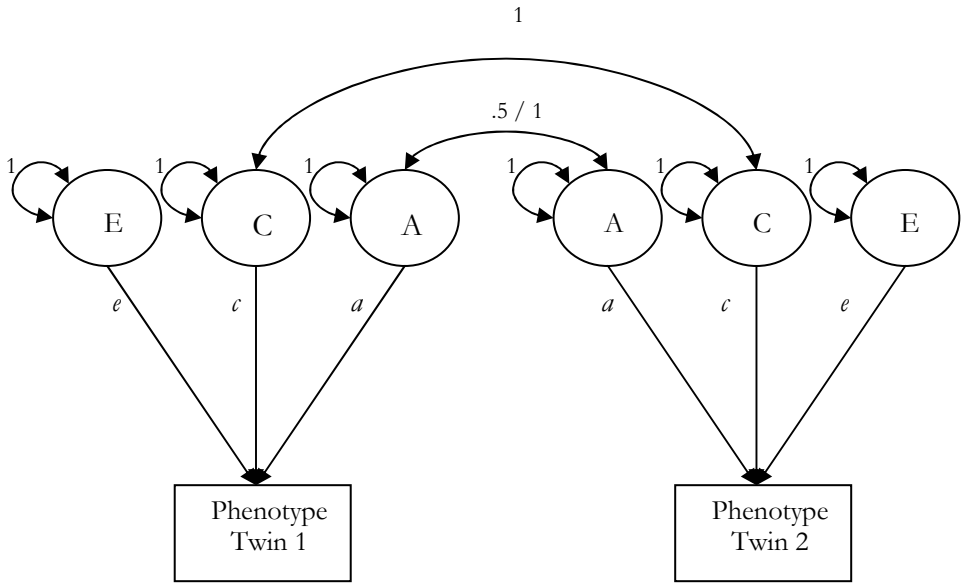


Figure 2.3. *A univariate path diagram*

As mentioned before (equation *2.20* and *2.21)*, the covariance for MZ and DZ twins are:

$$COV_{MZ} = a^2 + c^2 \qquad (eq.\ 2.23)$$
$$COV_{DZ} = \tfrac{1}{2}\, a^2 + c^2 \qquad (eq.\ 2.24)$$

A model according to the formulas of variance and twin covariance given above can be fitted to the data using structural equation modeling (SEM). In both univariate and multivariate genetic analysis, the identification of genetic and environmental parameters depends on a multigroup analysis in which data from MZ and DZ twins are analyzed simultaneously. The contribution of the latent variables are estimated as regression coefficients in the linear regression of the observed phenotype on the latent variables. In this method submodels are compared by hierarchic χ^2 tests. The χ^2 statistic is computed by subtracting –2(log-likelihood) for the full model from that for a reduced model ($\chi^2 = -2LL_0 - (-2LL_1)$). In addition to the χ^2 test statistic, Akaike's Information Criterion (AIC =

χ^2 – 2 x degrees of freedom) can be computed. The lower the AIC the better the fit of the model to the observed data.

Multivariate Designs

The advantage of structural equation modeling (e.g. in Mx, Neale *et al.*, 1999) is the expansion of relative simple univariate models to multivariate models were more than one measurement per subject is available. For instance, to investigate the correlation of two variable related to cognitive abilities (IQ and CITO) bivariate models can be used. Data collected repeatedly in time can be studied using longitudinal models. Further, variance in behavior assessed by different raters can be decomposed making use of multiple rater models. Rather than decomposing the variance of a measurement into genetic and environmental sources of variance, multivariate genetic analysis decomposes the variance of each measurement and the covariance between the measurements into genetic and environmental sources.

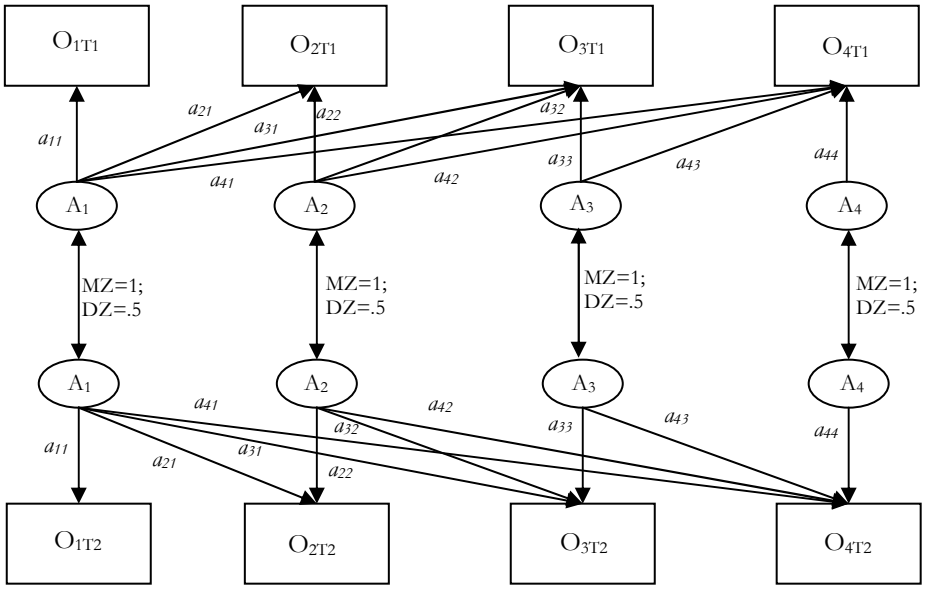


Figure 2.4.

A Cholesky Decomposition, in which O_1….O_4 represent the observed phenotype at measurement occasion one to measurement occasion four for twin 1 (T1) and twin 2 (T2) and A_1…A_4 represent the additive genetic latent factors. a_{11}..a_{44} represent the factor loadings. The influence of each variance component (A, C, or E) can be expressed this way.

The most commonly used model in multivariate designs is the saturated model, also known as a Cholesky Decomposition or a triangular decomposition (see Figure 2.4). Since the saturated model is fully parameterized, it yields the best possible fit to the data. The model is only descriptive and not driven by a specific hypothesis.

When data have been collected on different assessment points, the genetic and environmental influences can be estimated at each time interval separately. Using a longitudinal model, though, one can estimate how genes and environmental influences operate throughout development. The developmental mechanisms that were used are depicted in Figure 2.4, 2.5, and 2.6. We derive a matrix, $\sum$, representing the expected additive genetic covariance matrix. This matrix $\sum$ could, off course, also represent the covariance matrix for shared or nonshared environmental influences.

The saturated model in Figure 2.4, also known as a Cholesky or triangular decomposition, is an unconstrained model for the (co)variances among measurement occasions. It implies the covariance structure:

$$\sum = \mathbf{X}\,\mathbf{\Psi}\,\mathbf{X}' \qquad (eq.\ 2.25)$$

where $'$ indicates transposition. When the latent factors are independent, $\mathbf{\Psi} = \mathbf{I}$.
Matrix $\mathbf{X}$ is a $n_t \times n_t$ lower triangular matrix with n_t equal to the number of measurement occasions. For instance, for $n_t = 4$ matrix $\mathbf{X}$ would be:

$$\mathbf{X} = \begin{pmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

In the simplex model (Figure 2.5) there are 'carry-over' or transmission effects from one measurement occasion to the subsequent one as well as effects specific to each measurement occasion. This implies the covariance structure:

$$\sum = (\mathbf{I}-\mathbf{G})^{-1} \times (\mathbf{X}\,\mathbf{\Psi}\,\mathbf{X}') \times ((\mathbf{I}-\mathbf{G})^{-1})' \qquad (eq.\ 2.26)$$

Where $\mathbf{I}$ is a $n_t \times n_t$ identity matrix with elements on the main diagonal set to one. Matrix $\mathbf{G}$ is "sub" diagonal and contains the transmission effects ($\beta_{G(t)}$). If the innovation factors are independent $\mathbf{\Psi}=\mathbf{I}$.

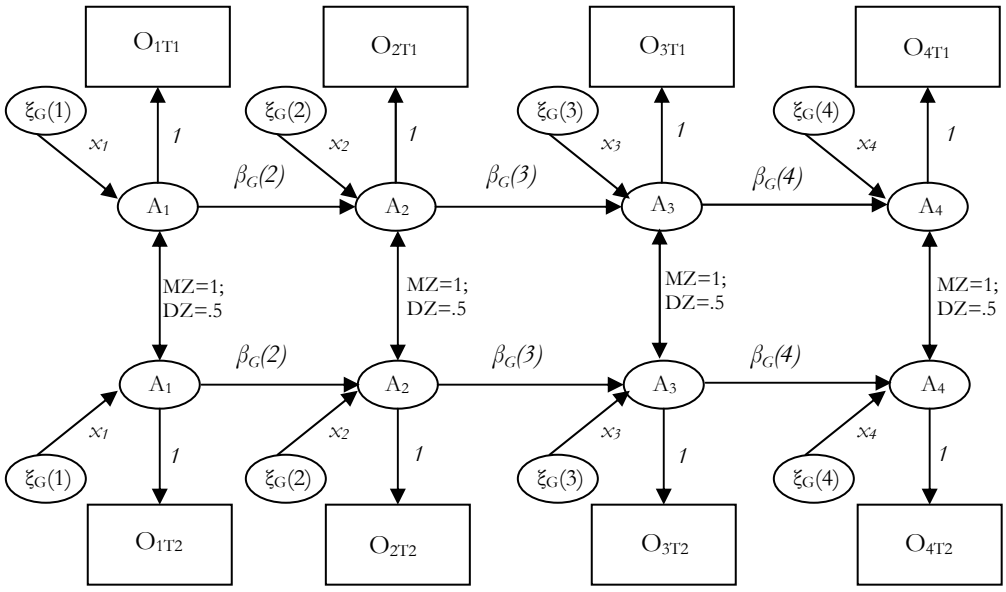


Figure 2.5.

A Simplex Model in which O_1….O_4 represent the observed phenotype at measurement occasion one to measurement occasion four for twin 1 (T1) and twin 2 (T2). $\xi_G(1)$… $\xi_G(4)$ represent the innovations and $\beta_G(2)$… $\beta_G(2)$ the transmission effects.

A first order autoregressive processes is assumed so that only the elements directly below the main diagonal are estimated. For instance, with four measurement occasions **G** would equal:

$$\mathbf{G} = \begin{pmatrix} 0 & & & \\ \beta_G 2 & 0 & & \\ 0 & \beta_G 3 & 0 & \\ 0 & 0 & \beta_G 4 & 0 \end{pmatrix}$$

Matrix **X** is a $n_t \times n_t$ diagonal matrix with parameters on the main diagonal and zeroes elsewhere. It represents the new influences ($\xi_G t$) that come into play at each measurement occasion.

$$\mathbf{X} = \begin{pmatrix} x_1 & & & \\ 0 & x_2 & & \\ 0 & 0 & x_3 & \\ 0 & 0 & 0 & x_4 \end{pmatrix}$$

In the common factor model depicted in Figure 2.6 one underlying factor (F_A) with time-specific factor loadings is specified. We assume one common factor in this project but the extension to multiple factors is straightforward. To account for occasion specific variance, time specific factors are added to the model. Assuming that the common and time specific factors are uncorrelated, the expected covariance matrix equals:

$$\sum = \mathbf{Q}\,\mathbf{\Psi}\,\mathbf{Q}' + \mathbf{X} \times \mathbf{X}' \qquad (eq.\ 2.27)$$

where $\mathbf{Q}$ is the $n_t \times 1$ vector with factor loadings and $\mathbf{X}$ is a $n_t \times n_t$ diagonal matrix containing the occasion-specific effects. $\mathbf{\Psi}=\mathbf{I}$, if the factors are independent.

$$\mathbf{Q} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{pmatrix} \quad \text{and} \quad \mathbf{X} = \begin{pmatrix} a_1 & & & \\ 0 & a_2 & & \\ 0 & 0 & a_3 & \\ 0 & 0 & 0 & a_4 \end{pmatrix}$$

The Cholesky, simplex, and factor model result in a different pattern of covariances among measurement occasions. The Cholesky is descriptive and merely estimates the variance-covariance matrix among time points. However, it is a useful model for evaluating the fit of more restricted models. For instance, if a factor model fits the data significantly poorer compared with the saturated model, it should be rejected. In a simplex model subsequent levels of the phenotype are influenced by prior levels. The implication of this autoregressive property is that effects of prior events or experiences will be larger to the extent that they happened closer in time (Guttman, 1954). The transmission model therefore predicts higher correlations among adjoining assessments than those occurring more distantly in time, the so-called simplex structure. In contrast, the factor model assumes that the same stable factors exert their effects at each assessment and does not imply that correlations between assessments vary as a function of

the length of the time lag. Thus, these models predict different patterns of longitudinal correlations, and tests can be performed to derive the underlying mechanism by comparing observed and predicted correlations.

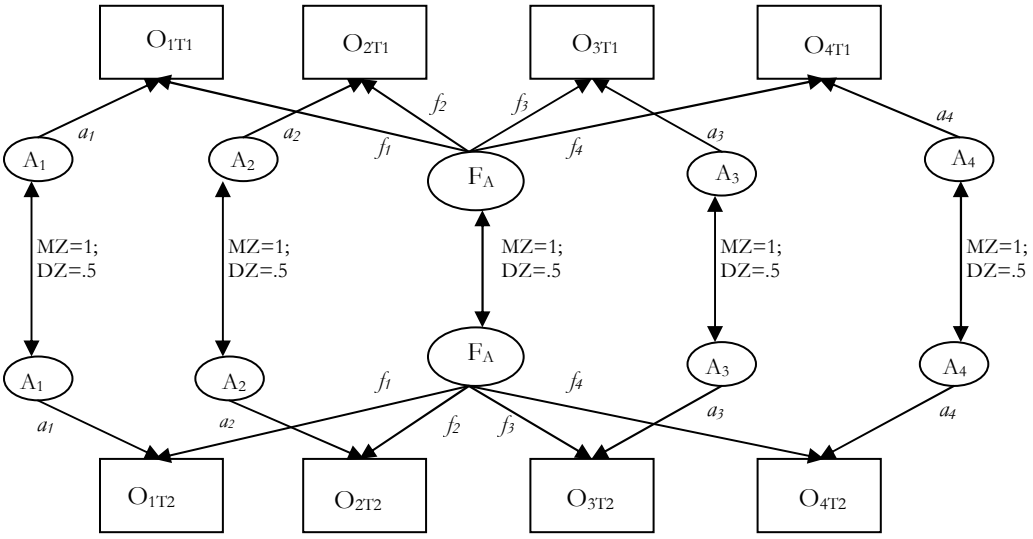


Figure 2.6.

A Common Factor Model in which O_1….O_4 represent the observed phenotype at measurement occasion one to measurement occasion four for twin 1 (T1) and twin 2 (T2). F_A *represent the common additive genetic factor. f_1…f_4 represent the factor loadings. A_1…A_4 represent the age specific influences.*

Multiple rater models

To investigate the underlying sources of parental disagreement two models can be considered. First, a so-called rater bias model (see Figure 2.7). As explained in the general introduction (Chapter 1), the Rater Bias model assumes that parents assess the same behaviors in the child and have a common understanding of the behavioral descriptions. This may apply when both parents are equally confronted with the behaviors shown by the child (for instance at home). Disagreement between the raters is regarded as error, resulting from rater bias and/or unreliability.

To derive the expected covariance matrix we first write this model for a single child in matrix form:

$$\begin{pmatrix} \mathrm{MRT}_1 \\ \mathrm{FRT}_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \times \begin{pmatrix} a \end{pmatrix} \times \begin{pmatrix} \mathrm{A} \end{pmatrix} + \begin{pmatrix} c \end{pmatrix} \times \begin{pmatrix} \mathrm{C} \end{pmatrix} + \begin{pmatrix} e \end{pmatrix} \times \begin{pmatrix} \mathrm{E} \end{pmatrix} + \begin{pmatrix} b_m & 0 \\ 0 & b_f \end{pmatrix} \times \begin{pmatrix} \mathrm{B_m} \\ \mathrm{B_f} \end{pmatrix} + \begin{pmatrix} r_m & 0 \\ 0 & r_f \end{pmatrix} \times \begin{pmatrix} \mathrm{R_m} \\ \mathrm{R_f} \end{pmatrix} \qquad (eq.\ 2.28)$$

When the latent factors are defined to have unit variance, the individual contributions of the latent factors can be obtained by taking the quadratic form of the parameter matrices is equation *2.28*:

$$\mathbf{A} = a^2,\ \mathbf{C} = c^2, \text{ and } \mathbf{E} = e^2$$

$$\mathbf{B} = \begin{pmatrix} b^2_m & 0 \\ 0 & b^2_f \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} r^2_m & 0 \\ 0 & r^2_f \end{pmatrix}$$

The expected covariance matrix of the Rater Bias model for two children of a twin pair will be:

$$\boldsymbol{\Sigma} = \mathbf{L} \times \begin{pmatrix} \mathbf{A} + \mathbf{C} + \mathbf{E} & r_g \otimes \mathbf{A} + \mathbf{C} \\ r_g \otimes \mathbf{A} + \mathbf{C} & \mathbf{A} + \mathbf{C} + \mathbf{E} \end{pmatrix} \times \mathbf{L}' + \begin{pmatrix} \mathbf{B} + \mathbf{R} & \mathbf{B} \\ \mathbf{B} & \mathbf{B} + \mathbf{R} \end{pmatrix} \qquad (eq.\ 2.29)$$

A, **C**, and **E** represent the additive genetic, shared environmental and non-shared environmental influences on the reliable trait variance (latent phenotype twin 1 and 2), representing behavior similar assessed by both parents. **B** represents the rater bias and **R** represents the residual variance or measurement error. ⊗ is the Kronecker product, **L** a matrix with loadings of additive genetic, shared environmental and nonshared environmental latent factors on the parental ratings, and r_g the correlation between the additive genetic factors of twins. The factor loading matrix is of the general form: $\mathbf{L} = \mathbf{I_t} \otimes (\mathbf{I_s} \otimes \mathbf{d})$ with $\mathbf{I_t}$ is a $n_t \times n_t$ identity matrix determined by the number of measurement occasions, $\mathbf{I_s}$ 2×2 identity matrix determined by the fact there are two children in a twin pair, and **d** is an $n_r \times 1$ vector determined by the number (n_r) of raters. Correlation r_g can be derived from quantitative genetic theory (Falconer and Mackay, 1996) and equals 1 for monozygotic twins and .5 for dizygotic twins.

In addition to assessing similar aspects of the child's behavior, the Psychometric model (Figure 2.8) assumes that each parent assesses a unique aspect of their child's behavior.

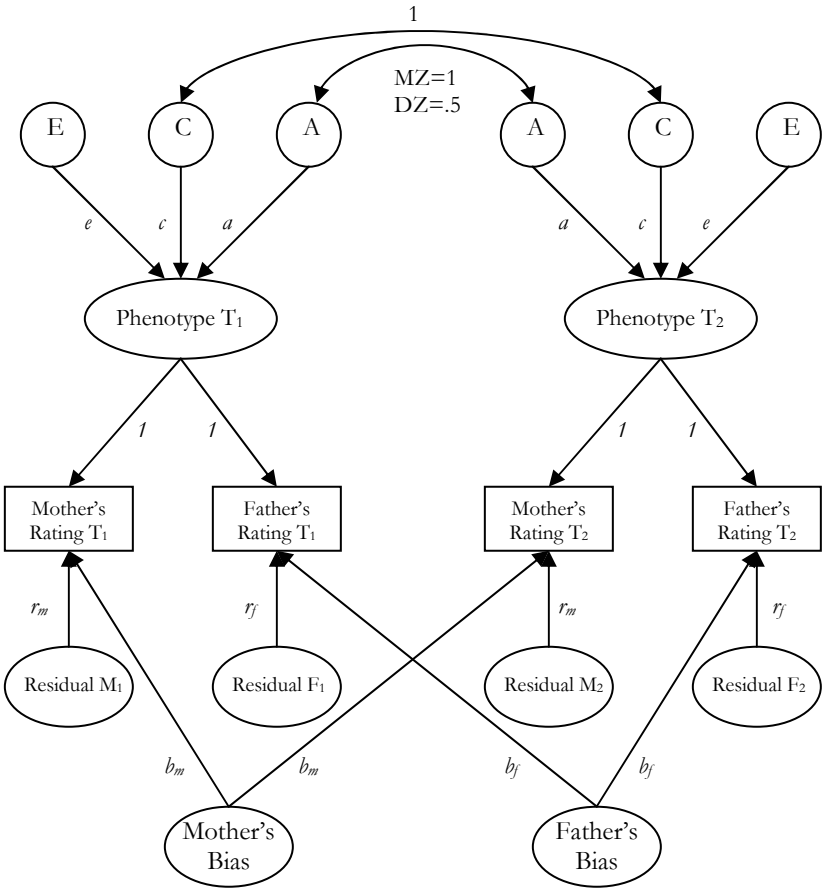


Figure 2.7. *Rater Bias model*

The psychometric model tests this possibility by examining whether there are significant genetic effects on the unique part of each parent's rating. If the behaviors uniquely rated by the parents are shown to be influenced by the genotype of the child, the parent must have been assessing a 'real' but unique aspect of the child's behavior. To derive the expected covariance matrix we first write this model for a single child in matrix form:

$$\begin{pmatrix} \mathrm{MRT}_1 \\ \mathrm{FRT}_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \times \begin{pmatrix} a \end{pmatrix} \times \begin{pmatrix} \mathrm{A} \end{pmatrix} + \begin{pmatrix} c \end{pmatrix} \times \begin{pmatrix} \mathrm{C} \end{pmatrix} + \begin{pmatrix} e \end{pmatrix} \times \begin{pmatrix} \mathrm{E} \end{pmatrix} + \begin{pmatrix} a_m & 0 \\ 0 & a_f \end{pmatrix} \times \begin{pmatrix} \mathrm{A_m} \\ \mathrm{A_f} \end{pmatrix} + \begin{pmatrix} c_m & 0 \\ 0 & c_f \end{pmatrix} \times \begin{pmatrix} \mathrm{C_m} \\ \mathrm{C_f} \end{pmatrix} + \begin{pmatrix} e_m & 0 \\ 0 & e_f \end{pmatrix} \times \begin{pmatrix} \mathrm{E_m} \\ \mathrm{E_f} \end{pmatrix} \quad (eq.\ 2.30)$$

When the latent factors are defined to have unit variance, the individual contributions of the latent factors can be obtained by taking the quadratic form of the parameter matrices is equation *2.30*:

$$\mathbf{A} = a^2, \ \mathbf{C} = c^2, \text{ and } \mathbf{E} = e^2$$

$$\mathbf{G} = \begin{pmatrix} a^2_m & 0 \\ 0 & a^2_f \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} c^2_m & 0 \\ 0 & c^2_f \end{pmatrix}, \text{ and } \mathbf{F} = \begin{pmatrix} e^2_m & 0 \\ 0 & e^2_f \end{pmatrix}$$

The expected covariance matrix of the Psychometric model for two children of a twin pair will be:

$$\boldsymbol{\Sigma} = \mathbf{L} \times \begin{pmatrix} \mathbf{A} + \mathbf{C} + \mathbf{E} & r_g \otimes \mathbf{A} + \mathbf{C} \\ r_g \otimes \mathbf{A} + \mathbf{C} & \mathbf{A} + \mathbf{C} + \mathbf{E} \end{pmatrix} \times \mathbf{L}' + \begin{pmatrix} \mathbf{G} + \mathbf{S} + \mathbf{F} & r_g \otimes \mathbf{G} + \mathbf{S} \\ r_g \otimes \mathbf{G} + \mathbf{S} & \mathbf{G} + \mathbf{S} + \mathbf{F} \end{pmatrix} \quad (eq.\ 2.31)$$

where, as mentioned above, ⊗ is the Kronecker product, **L** a matrix with loadings of the common factors on the parental ratings, and r_g the correlation between the additive genetic factors of twins. **A**, **C**, and **E** represent the additive genetic, shared environmental and non-shared environmental influences on the reliable trait variance, representing behavior similar assessed by both parents. **G** is the unique additive genetic variance, representing the parental unique view on the child's behavior. **S** is the unique shared environmental factor, representing rater bias. Finally, **F** is the unique nonshared environmental factor, representing measurement error.

In multivariate structural equation modeling the longitudinal and multiple rater models can be combined to exhibits the advantages of genetically informative data. It can be shown that these data are not merely useful for estimating the size of genetic and environmental effects but have the potential to shed light on fundamental questions and would therefore be a useful addition to the traditional method arsenal for studying psychological data.

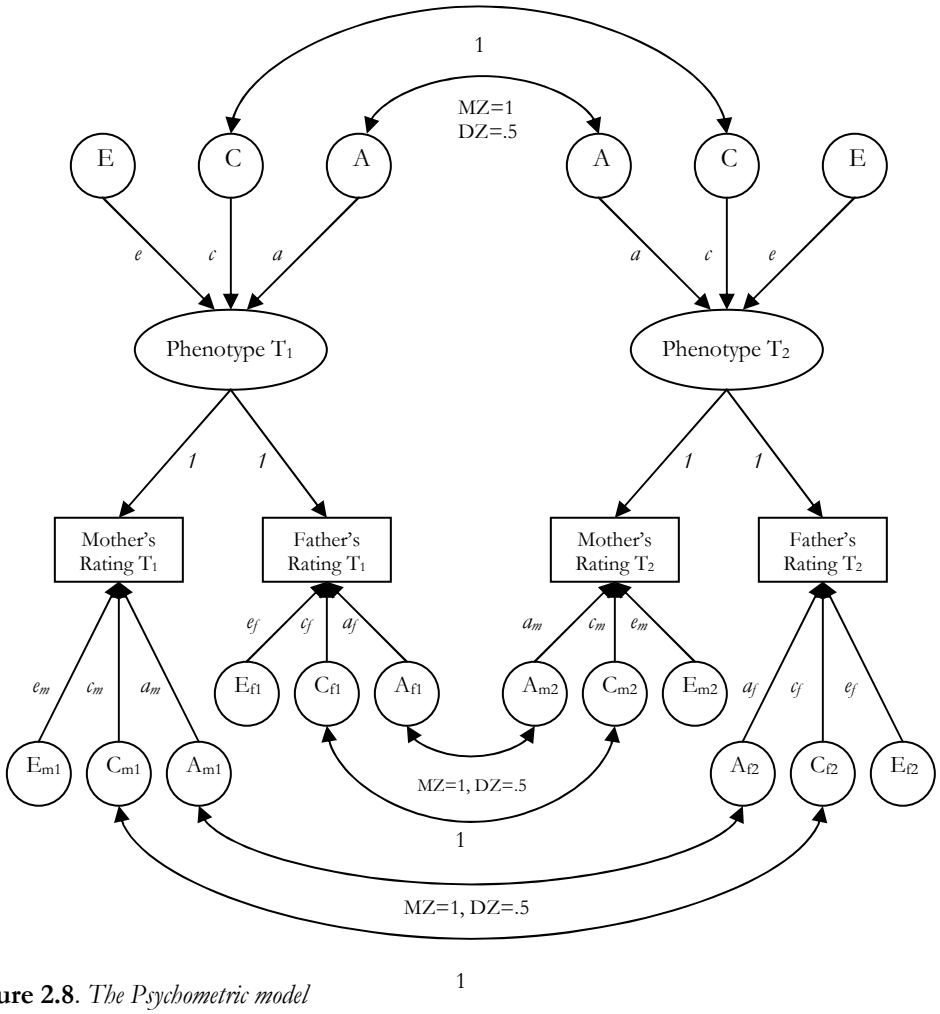


Figure 2.8. *The Psychometric model*

3

Genetic and environmental mechanisms underlying stability and change in problem behaviors at the ages 3, 7, 10, and 12

This chapter is submitted as:

Bartels, M., Van den Oord, E.J.C.G., Hudziak, J.J., Rietveld, M.J.H., van Beijsterveldt, C.E.M., and Boomsma, D.I. (2002). Genetic and environmental mechanisms underlying stability and change in problem behaviors at ages 3, 7, 10, and 12. *Developmental Psychology*, submitted.

ABSTRACT

We performed an unique longitudinal genetic study of behavior problems in a large sample of Dutch twins. The number of participating twin pairs at ages 3, 7, 10 and 12 were 5,602, 5,115, 2,956, and 1,481, respectively. Genetic, shared environmental, and nonshared environmental factors accounted for 43/60, 47/34, and 10/6 percent of the stability in Internalizing/Externalizing problems. The genetic contribution resulted from the fact that a subset of genes expressed at an earlier age was still active at the next time point. A common set of shared environmental factors seemed to operate at all ages. The modest contribution of nonshared environmental factors could not be captured by a simple model. Significant age specific influences were found for all components. This indicated that genetic and environmental factors also contributed to changes in problem behaviors.

INTRODUCTION

Several studies have reported a substantial degree of continuity in problem behaviors. For example, Richman and colleagues (1982) found that 61% of the problematic children at age 3 still showed considerable difficulties on a clinical rating scale five years later. Graham and Rutter (1973) showed that 75% of the 10-11-year-old children who received a diagnosis of conduct disorder and 46% of the children, who received a diagnosis of emotional disorder, remained deviant at the follow up four years later. Stability is not confined to clinical groups and has also been found in general population samples. Verhulst and Van der Ende (1992b) reported a correlation of .56 for problem behaviors across a six-year period in a population sample of Dutch children originally aged 4-11 years. Ghodsian and colleagues (1980) studied a national sample of British children assessed at ages 7, 11, and 16. Correlations for parental ratings of problem behavior were .48 between ages 7 and 11, .38 between 7 and 16 years, and .46 between 11 and 16 years. Although these studies indicated a substantial degree of continuity, problem behaviors should not be regarded as static (Verhulst and Van der Ende, 1992a). Many children also show changes across time. For a comprehensive picture of developmental processes underlying child psychopathology the mechanisms that explain continuity and change need to be understood. Genetic and environmental influences, for instance, can yield similar or distinct influences on the developmental process.

A powerful tool to unravel the genetic and environmental architecture of individual differences in the development of behavioral and emotional problems is to study genetically related individuals. Family studies might give a first impression of familial aggregation, but they cannot distinguish between genetic and environmental effects. Similarities between family members may be created either by genetic relatedness or by sharing the same family environment. A method, which solves this problem, is the classical twin design. Monozygotic (MZ) twins derive from a single zygote and therefore two individuals of a MZ twin pair are genetically identical. Dizygotic (DZ) twins develop from two distinct zygotes and share on average 50% of their genes, like ordinary brothers and sisters. Hence, the only possible way to explain the variation in problem behavior between two members of a MZ twin pair are environmental effects that are not shared by those two: the so-called nonshared environmental influences. Conversely, the variation in problem behavior between two members of a DZ twin pair could result from different genes and/or nonshared environmental influences. Accordingly, the difference in relatedness between MZ and DZ twin pairs (mostly expressed as correlation coefficients: r_{MZ} and r_{DZ}) gives information about the strength of the genetic and environmental influences on the trait under investigation.

Longitudinal twin and family data allow the study of persistence and change of genetic (A), shared environmental (C), and nonshared environmental (E) influences. First, genetic or environmental factors may exert a continuous influence from their time of onset (common factor influences). This mechanism implies that the same genetic or environmental factors are responsible for stability, possibly with age-dependent factor loadings. Second, there can be a simplex-like continuity in genetic and environmental effects (Eaves *et al.*, 1986; Boomsma and Molenaar, 1987). In this simplex-like continuity, there are effects specific to each age and there are 'carry-over effects' or transmission effects from one age to the subsequent age. In other words, earlier influences may be transmitted from one occasion to the next and new influences (innovations) may come into play at each occasion.

Simplex and factor models both imply a certain degree of continuity. However, in a longitudinal study, these mechanisms result in a different pattern of correlations between successive assessments. In a simplex model subsequent levels of problem behavior are influenced by prior levels. The implication of this autoregressive property is that effects of prior events or experiences will be larger to the extent that they happened closer in time (Guttman, 1954). The transmission model therefore predicts higher correlations among adjoining assessments than those occurring more distantly in time, the so-called simplex structure. In contrast, the factor model assumes that the same stable factors exert their effects at each assessment and does not imply that correlations between assessments vary as a function of the length of the time lag. Thus, simplex and factor models predict different patterns of longitudinal correlations, and tests can be performed to derive the underlying mechanism by comparing observed and predicted correlations.

Only a few studies used this method to disentangle the genetic an environmental influences on continuity and change in the development of problem behaviors or problem behavior related disorders. In three studies the Child Behavior Checklist (CBCL; Achenbach, 1991; Achenbach, 1992) was used to rate problem behavior in children. First, Van der Valk and colleagues (2002) used a two-wave behavior genetic model to estimate genetic, shared environmental and nonshared environmental contributions to stability and change of Internalizing and Externalizing Problems at ages 3 and 7 years in a Dutch sample overlapping to our sample. For Externalizing problems the estimated influences of additive genetic, shared and nonshared environmental factors remained relatively constant over the years. The phenotypic stability (r=.54) was explained for 55% by genetic factors. Shared environmental influences were mostly stable, while nonshared environmental influences were mostly age specific. For Internalizing problems additive genetic influences decreased while nonshared environmental influences increased over the years. The phenotypic stability (r=.38) was for 66% explained by genetic factors. Second, Schmitz *et*

al. (1995) conducted a study, examining a small longitudinal sample of 95 twin pairs, assessed at ages 2 and 7 years. Results indicated that for Internalizing Problems continuing shared environmental factors had an effect both in early and middle childhood, while genetic influences had mostly age specific effects. For Externalizing Problems the opposite effect was found, showing continuing genetic and age specific shared environmental effects. However, as suggested by the authors, these results need to be replicated in larger samples of genetically informative data. A study with biologically related and unrelated adoptees suggested that stability in Externalizing problem behavior over a three-year interval is mainly accounted for additive genetic influences while nonshared environmental influences mainly account for stability in Internalizing problem behavior (Van der Valk *et al.*, 1998b)

A developmental study in sibs, half sibs and cousins, by Van den Oord and Rowe (1997), looked at maternal ratings of The Behavior Problems Index (Peterson and Zill, 1986). There were 436 pairs of full siblings, 119 pairs of half siblings, and 122 pairs of cousins assessed at ages 4-6, 6-8, and 8-10. In this study, the continuity of problem behaviors was entirely explained by genetic and shared environmental factors. Nonshared environmental factors only showed age specific effects, influencing changes in children's problem behaviors. O'Connor *et al.* (1998), followed 405 families over a three-year interval. Subjects consisted of monozygotic and dizygotic twins, and full, half and unrelated siblings (all of same-sex) between 10 and 18 years of age at the first assessment. Results showed that the phenotypic stability of antisocial symptoms of $r = .63$ was explained for 54% by continuing genetic influences and for 30% by continuing shared environmental influences. For depressive symptoms, the phenotypic stability of $r = .59$ was explained for 64% by continuing genetic influences and for 36% by continuing nonshared environmental influences. In short, even though each study investigated subjects at a different age interval, most studies showed large influences for genetic factors on the stability of problem behaviors. Effects of shared and nonshared environmental factors are less clear, showing continuing influences for some studies and only age specific effects for others.

In the present longitudinal study structural equation modeling techniques were used to examine the influences of genetic and environmental factors on development of Internalizing and Externalizing behavior, using data of a large sample of Dutch twin pairs at 3, 7, 10, and 12 years of age. In addition to estimating the importance of heritability and environmental influences, the focus was on the mechanism underlying the developmental pattern of behavior. A genetic simplex model and a common factor model were used to study continuity and changes of genetic and environmental influences over time.

METHODS

Subjects

All participants were registered by the Netherlands Twin Registry (NTR), kept by the Department of Biological Psychology at the Vrije Universiteit in Amsterdam. Of all multiple births in the Netherlands, 40-50% is registered by the NTR (Boomsma *et al.*, 1992; Boomsma, 1998; Boomsma *et al.*, 2002). For this study, data from twins from the birth cohorts 1986 - 1993 were used. Behavioral questionnaires have been collected longitudinally at ages 3, 7, 10, and 12. Child Behavior Checklists (CBCL; Achenbach, 1991; Achenbach, 1992) were mailed to families within three months of the twins' third, seventh, tenth, and twelfth birthday. Two to three months after the mailing reminders were sent and four months after the initial mailing persistent non-responders were contacted by phone. Families whose address was no longer available were included in the nonresponse group. From the original sample 253 twin pairs were excluded because either one or both of the children had a disease or handicap that interfered severely with daily functioning at 12 or at a younger age. Thus, at age 3 maternal ratings were available for 5,602 twin pairs, at age 7 maternal ratings were available for 5,115 twin pairs, at age 10 maternal ratings were available for 2,956 twin pairs and at age 12 maternal ratings were available for 1,481 twin pairs.

Zygosity was determined for 787 same-sex twin pairs by DNA analyses or blood group polymorphisms. For all other same-sex twin pairs zygosity was determined by discriminant analysis, using questionnaire items at each age separately. Parents were asked how much the twins resembled each other in facial structure, hair color, facial color, eye color, and whether they were ever mistaken for each other by the parents themselves, by family, or by strangers. They were also asked if the twins were as much alike as two peas in a pod, whether it was difficult for the parents to separate the twins on a recent picture, and how similar the twins' hair structure was. Agreement between zygosity assignment by the replies to the questions and zygosity determined by DNA markers/blood typing is around 93% (For details see Rietveld *et al.*, 2000).

Measures

Mother ratings were collected by making use of the Child Behavior Checklist (Achenbach, 1991; Achenbach, 1992). The Checklist for 2-3-year-old children (CBCL 2/3) shows age-adjusted differences from the checklist for 4 to 18-year old children (CBCL 4-18).

The CBCL 2/3 was developed for parents to score the behavioral and emotional problems of their 2- and 3-year-old children. It consists of 100 items that are scored by the parents on a 3-point scale based on the occurrence of the behavior during the preceding 2 months: 0 if the problem item was not true, 1 if the item was somewhat or sometimes

true, and 2 if it was very true or often true. Dutch syndrome scales and comparability with the syndrome scales as developed by Achenbach (1992) are reported by Koot and colleagues (1997). In the present paper the two broad band scales Internalizing and Externalizing are analyzed. The Internalizing scale consists of the Anxious and Withdrawn/Depressed subscales. The Externalizing scale consists of the Aggressive, Oppositional, and Overactive subscales. For the Internalizing scale subjects were only included if not more than 1 item was missing for the Anxious, and not more than 2 items were missing for the Withdrawn/Depressed scale. For the Externalizing scale the inclusion criterion was not more than 1 item missing for the Aggressive and the Overactive and not more than 3 items for the Oppositional scale. This ensured that the two syndrome scales were always composed of all problem behaviors loading on that scale.

The CBCL 4-18 was developed for parents to score the behavioral and emotional problems of their 4-to-18-year-old children. It consists of 120 problem items that are scored by the parents on a 3-point scale based on the occurrence of the behavior during the preceding 6 months: 0 if the problem item was not true, 1 if the item was somewhat or sometimes true, and 2 if it was very true or often true. The syndrome scales were composed according to the 1991 profile (Achenbach, 1991). Dutch syndrome scales and comparability with the syndrome scales as developed by Achenbach are reported in Verhulst *et al.* (1996). In this manual the two broad-band scales Internalizing and Externalizing are analyzed. The Internalizing scale consists of the Anxious/Depressed, Somatic Complaints and Withdrawn subscales. The Externalizing scale consists of the Aggressive and Rule Breaking Behavior subscales. For the Internalizing scale subjects were only included if not more than 3 items were missing for the Anxious/Depressed scale, not more than 2 items were missing for Somatic Complaints and Withdrawn scales. For the Externalizing scale the inclusion criterion was not more than 3 items missing for the Aggressive and Rule Breaking Behavior scales. This ensured that the two syndrome scales were always composed of all problem behaviors loading on that scale.

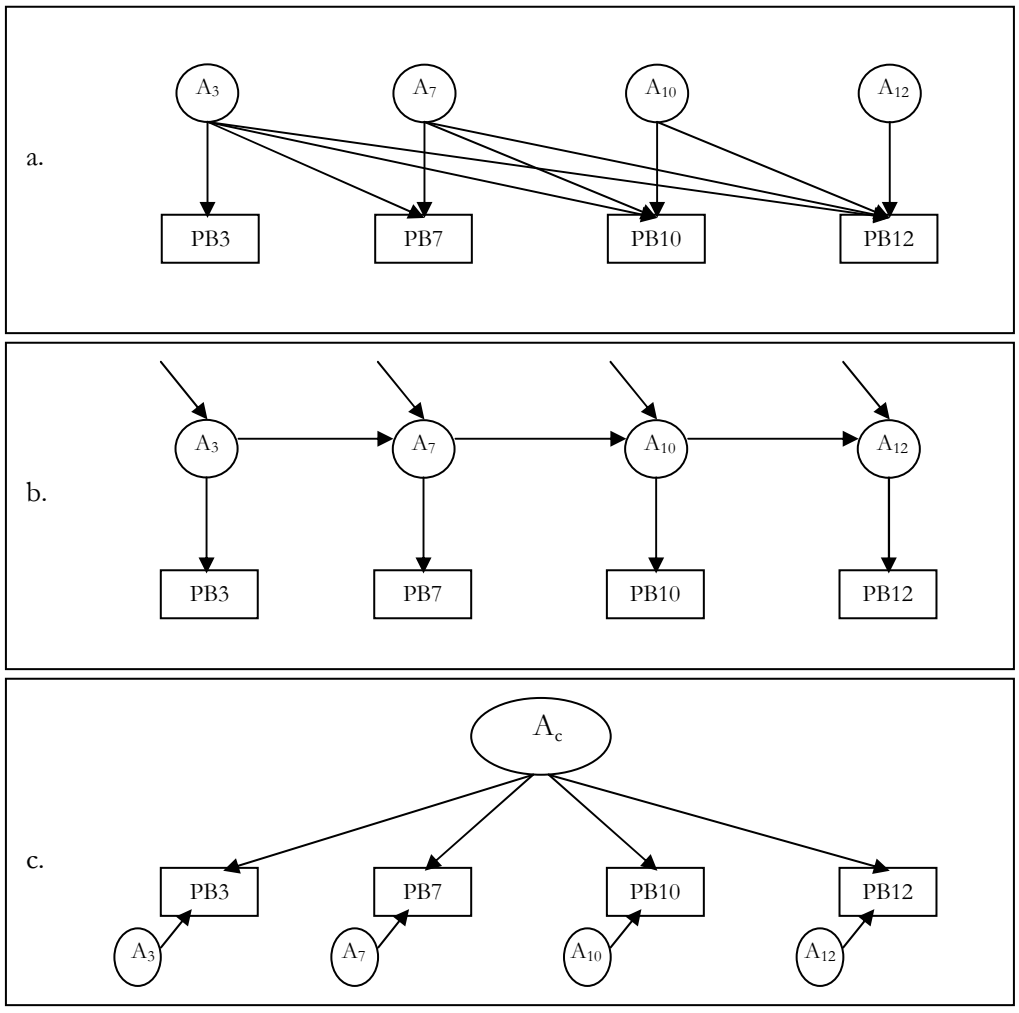
Genetic modeling

Genetic model fitting of twin data allows for decomposition of the observed phenotypic variance into its genetic and environmental components. Additive genetic variance (A), is the variance that results from the additive effects of alleles at each contributing genetic locus. Shared environmental variance (C) is the variance that results from environmental events common to both members of a twin pair. Nonshared environmental variance (E) is the variance that results from environmental effects that are not shared by members of a

twin pair. Estimates of the nonshared environmental effects also include measurement error. To account for this source of variance, E is always specified in the model.

Figure 3.1.

The three models used to investigate the underlying process of the development of problem behavior (PB3, PB7, PB10, PB12). a. the saturated model; b. the simplex model; c. the common factor model, with age specific influences Obviously, all three variance components (A, C, E) can be expressed in either way.



The different degree of genetic relatedness between monozygotic (MZ) and dizygotic (DZ) twin pairs was used to estimate the contribution of these factors to the phenotypic variation in problem behavior. Similarities for MZ twins are assumed to be due to additive

genetic influences plus environmental influences that are shared by both members of a twin pair. Experiences that make MZ twins different from one another are nonshared environmental influences. Because DZ twins share 50% of their genetic material on average, like full siblings, genetic factors contribute only half to their resemblance. As for MZ twins the shared environment contributes fully. Model fitting to twin data is based on the comparison of the variance-covariance matrices in MZ and DZ twins. The whole variance-covariance matrix can be decomposed into a matrix of genetic variances and covariances, a matrix of shared environmental variances and covariances, and a matrix of nonshared environmental variances and covariances.

Multivariate genetic model fitting techniques were used to obtain insight in the developmental pattern of Internalizing and Externalizing problem behavior and to obtain estimates of the genetic and environmental influences on variances and covariances in problem behavior. In the present study three types of models were used (See Figure 3.1): saturated models (3.1a), simplex models (3.1b) and common factor models (3.1c). For each of the models, the total variances and covariances were decomposed into additive genetic (A), shared environmental (C), and nonshared environmental (E) parts. The saturated model, also know as a Cholesky Decomposition or triangular decomposition, decomposes the phenotypic statistics into genetic, shared environmental, and nonshared environmental contributions. Since the saturated model is fully parameterized, it yields the best possible fit to the input matrices. The model is only descriptive and not driven by a specific hypothesis. However, it is a useful model for evaluating the fits of more restricted models. This evaluation consists of performing tests that compare the genetic and environmental contributions as predicted by a developmental model with the unconstrained genetic and environmental contributions from the saturated model. If a developmental model fits the data significantly poorer compared with the saturated model, the predicted contributions are inconsistent with the data, and the developmental model should be rejected.

The simplex model (Figure 3.1b) is a first order auto-regressive process. In the simplex model covariances among the four ages of measurement are specified by genetic and environmental factors specific to each age and by 'carry-over effects' or transmission of these factors to subsequent ages. The model specifies the variance unique to each measurement occasion by an innovation term that comes into play at each time point. The total variance is the sum of the age specific effects and age-to-age transmission effects.

In a common factor model (Figure 3.1c) one underlying factor with age-specific factor loadings is specified, which implies a continuous influence from time of onset. To account for some age-specific variance, age-specific influences are added to the model.

The order of model reduction and the possibilities of model specification influence the results of the parameter estimates and the goodness of fit procedure. To take this into account all three variance components A, C, or E were analyzed separately, leaving the other two expressed in a saturated model. Finally, a 'best' model was fitted to the data to obtain estimates of additive genetic, shared environmental and nonshared environmental influences on the variances and covariances of Internalizing and Externalizing behavior.

Since in a longitudinal design data from one or more measurement occasions or from one twin may be missing from the dataset, multivariate structural models were fit to the transformed raw data by the method of maximum likelihood pedigree analysis (Lange *et al.*, 1976) using the statistical software package Mx (Neale *et al.*, 1999). Parameter estimates including those for means of each variable for the first and second twin of each zygosity group were produced that maximizes the joint likelihood of the raw data under a given structural model (Neale *et al.*, 1999; Neale and Cardon, 1992). In order to use this method, the data were square-root transformed to approximate normal distributions that are required for maximum likelihood estimation. After transformation, all skewness and kurtosis indices were between -1.0 and 1.0, implying that not much distortion is to be expected (Muthén and Kaplan, 1985).

Submodels were compared by hierarchic χ^2 tests. The χ^2 statistic is computed by subtracting –2(log-likelihood) for the full model from that for a reduced model ($\chi^2 = (-2LL_0 - (-2LL_1))$). In addition to the χ^2 test statistic, Akaike's Information Criterion (AIC = χ^2 – 2 x degrees of freedom) was computed. The lower the AIC the better the fit of the model to the observed data. Although the simplex model and the factor model do not form a nested pair, they may be compared in terms of parsimony and goodness of fit because they represent alternative sets of constraints on the saturated model (Neale and Cardon, 1992). Fit statistics of the reduced developmental models are compared to the saturated models. This results in a χ^2 and AIC, which are comparable for the different reduced models.

RESULTS

The untransformed mean problem scores and standard deviations of the twin sample and those of comparable community samples (Koot, 1997 and Verhulst *et al.*, 1996) are given in Table 3.1. For both the Internalizing and the Externalizing scale, the ratings given to the twins were quite similar to the ratings given to the Dutch community sample.

Table 3.1

Means (standard deviation) and sample sizes for the Internalizing and Externalizing scales in 3, 7, 10, and 12-year-old twins (per zygosity) compared to a 2- and 3-year-old Dutch community sample and a 4-11-year-old Dutch community sample

			INTERNALIZING		EXTERNALIZING	
			Mother	**N**	**Mother**	**N**
3	♂	**MZM**	4.64 (4.00)	1740	17.94 (10.49)	1738
		DZM	4.59 (4.05)	1752	16.92 (9.90)	1752
		DOS	4.52 (3.82)	1771	16.04 (9.83)	1769
		COM	4.5 (4.4)	215	17.5 (9.5)	215
	♀	**MZF**	4.78 (4.08)	2038	15.64 (10.01)	2035
		DZF	4.90 (4.18)	1637	15.30 (9.71)	1641
		DOS	3.98 (3.77)	1770	14.26 (9.43)	1772
		COM	4.3 (3.6)	205	16.5 (8.8)	205
7	♂	**MZM**	4.31 (4.11)	1719	9.68 (7.24)	1735
		DZM	4.78 (4.74)	1616	8.96 (7.22)	1650
		DOS	4.24 (4.44)	1560	8.45 (7.02)	1595
		COM	4.52 (4.27)	579	8.26 (6.36)	579
	♀	**MZF**	5.05 (4.77)	1965	6.90 (6.30)	1985
		DZF	5.21 (4.91)	1577	6.77 (6.02)	1611
		DOS	4.52 (4.31)	1570	6.25 (6.09)	1596
		COM	5.16 (5.02)	593	6.04 (5.57)	593
10	♂	**MZM**	4.59 (4.60)	1008	8.67 (7.43)	1029
		DZM	5.18 (5.44)	933	8.03 (7.34)	937
		DOS	4.66 (4.97)	877	7.69 (7.10)	885
		COM	4.52 (4.27)	579	8.26 (6.36)	579
	♀	**MZF**	5.12 (5.00)	1216	5.91 (5.53)	1223
		DZF	5.35 (5.35)	893	5.90 (5.86)	905
		DOS	4.76 (4.91)	878	5.34 (5.37)	885
		COM	5.16 (5.02)	593	6.04 (5.57)	593
12	♂	**MZM**	4.02 (4.53)	552	7.20 (6.90)	557
		DZM	4.05 (4.61)	446	6.67 (6.47)	460
		DOS	3.79 (4.68)	410	6.70 (6.76)	416
		COM	5.36 (5.36)	440	6.35 (6.13)	440
	♀	**MZF**	4.72 (4.70)	612	5.17 (5.27)	624
		DZF	4.56 (4.42)	440	4.86 (5.06)	451
		DOS	4.17 (4.31)	405	4.75 (5.31)	418
		COM	6.32 (5.93)	456	5.21 (5.43)	456

In previous studies, using an overlapping sample, comparable levels of problem behavior were found (Van der Valk *et al.*, 1998a, 2001, 2002: in press; Bartels et al, 2002a, submitted). Table 3.2 shows the phenotypic correlations (upper part of table), twin correlations (lower part of table, diagonal), and the twin cross correlations (lower part of table, off-diagonal) for Internalizing and Externalizing problem behavior at ages 3, 7, 10, and 12.

Table 3.2.

Phenotypic correlations, twin correlations and cross correlations for Internalizing (above diagonal) and Externalizing (below diagonal) problem behavior.

	3	**7**	**10**	**12**
3	1	.37 (.35-.37)	.33 (.31-.33)	.30 (.30-.33)
7	.55 (.54-.55)	1	.62 (.62-.63)	.57 (.54-.59)
10	.49 (.49-.51)	.73 (.71-.73)	1	.67 (.67-.67)
12	.48 (.48-.50)	.68 (.67-.70)	.75 (.75-.77)	1
MZ	**3**	**7**	**10**	**12**
3	.72 (.72-.74)\.84 (.82-.85)	.35 (.31-.35)	.31 (.27-.35)	.29 (.24-.35)
7	.54 (.50-.57)	.72 (.70-.74)\.86 (.85-.87)	.52 (.49-.56)	.48 (.43-.52)
10	.48 (.44-.51)	.67 (.64-.69)	.67 (.51-.70)\.84 (.82-.86)	.52 (.48-.57)
12	.47 (.42-.52)	.63 (.63-.67)	.69 (.66-.72)	.73 (.69-.76)\.86 (.84-.88)
DZ	**3**	**7**	**10**	**12**
3	.40 (.37-.42)\.55 (.53-.58)	.26 (.24-.29)	.28 (.24-.31)	.26 (.21-.30)
7	.35 (.33-.38)	.50 (.47-.53)\.55 (.53-.58)	.39 (.36-.42)	.35 (.31-.39)
10	.33 (.30-.36)	.42 (.39-.45)	.51 (.48-.55)\.52 (.49-.55)	.38 (.35-.43)
12	.33 (.29-.36)	.40 (.40-.44)	.43 (.39-.46)	.54 (.49-.59)\.57 (.52-.61)

The phenotypic correlations presented in Table 3.2 give a first impression of the underlying developmental pattern of Internalizing and Externalizing Problem Behavior. For both Internalizing and Externalizing behavior the phenotypic cross correlations are lower for longer intervals. This structure suggests a simplex pattern. This simplex structure, though, cannot be the sole mechanism because the stability between the distinct ages is higher than would be expected based on a simplex structure. For instance, if a simplex structure is describing the developmental process of Internalizing problem behavior, the product of the cross correlations between age 7 and 10 and age 10 and 12 is about .42 (.62 x .67). However, the real cross correlation between age 7 and 12 is .57. This

higher than expected cross-correlation implies, besides transmission, the influence of an underlying common factor on the development of childhood psychopathology. The lower part of Table 3.2 presents twin correlations (diagonal) and cross-correlations (off-diagonal). Cross correlations represent, for instance, the correlation between Internalizing behavior in the oldest twin at age 3 and Internalizing behavior in the youngest twin at age 7. In other words, this cross correlations is a cross-twin-cross-age correlation, which gives information on the development of problem behavior over the years. For both Internalizing and Externalizing problem behavior MZ twin correlations are higher than DZ twin correlations suggesting genetic influences at each age. The MZ correlations, though, are less than twice the DZ correlations indicating significant influences of shared environment as well. Cross-sectional estimates for additive genetic, shared and nonshared environmental influences for both Internalizing and Externalizing problem behavior based on the univariate model-fitting procedure can be found elsewhere (Van der Valk *et al.*, 1998a, 2001; Van der Valk *et al.*, in press; Bartels *et al.*, 2002a submitted; Bartels *et al.*, 2002d, submitted). To get insight in the presence or absence of sex differences twin correlations and cross-correlations for each zygosity group (MZM, DZM, MZF, DZF, DOS) separately have been calculated (Appendix 3.1). No heterogeneity is expected based on the fact that the correlations in DZ twins of opposite sex are not lower than the correlations in same sex DZ twins.

Cross-correlations for MZ and DZ twins were calculated to explore the genetic and environmental influences on the observed stability. As can be seen in Table 3.2, the MZ cross correlations are higher than the DZ cross correlations, but certainly not twice has high, suggesting that stability in Internalizing and Externalizing problem behavior over time is due to additive genetic factors as well as shared environmental factors.

Model-fitting procedures for Internalizing and Externalizing problem behavior yielded the results presented in Table 3.3. The saturated, Cholesky Decomposition, model without restrictions (model 1) was taken as a reference for evaluating changes in χ^2 and associated degrees of freedom of more parsimonious models. Because of the fact that the order of model reductions could influence the results of the goodness of fit procedure all three variance components A, C, and E were analyzed separately. The mechanism of one variance component at the time was investigated leaving the other two expressed in a saturated model (model 2 to model 7). Model 8 and model 9 represent the best fitting simplified models with and without sex differences. Sex differences in the strength of the genetic and environmental effects are tested by constraining the influences for boys and girls to be equal.

For Internalizing problem behavior a best fitting simplified model to describe the processes of development is a model with a simplex structure for additive genetic factors and a factor structure with age specific factor loadings for shared environmental influences (model 9). As can be seen in Table 3.3, for additive genetic influences the factor (model 3) and the simplex (model 2) model gave almost identical fits to the data. The simplex model, however, is a more parsimonious model and therefore preferred over a factor model. This parsimony is reflected in a lower AIC for the simplex model than for the factor model, suggesting a better fit of the simplex model. No sex-differences are found for Internalizing problem behavior.

Table 3.3.

Multivariate model fitting for Internalizing problem behavior at age 3,7,10 and 12

INTERNALIZING							
MODEL	**-2LL**	**df**	**c.t.m.**[a]	**χ^2**	**df**	**p**	**AIC**
1. Saturated model **A**: Saturated **C**: Saturated **E**: Saturated	77086.981	29285					
2. **A**: Simplex Structure	77095.034	29291	1	8.053	6	.23	-3.95
3. **A**: Factor Structure	77094.903	29286	1	7.922	4	.09	-.08
4. **C**: Simplex Structure	77102.743	29291	1	15.762	6	.02	3.76
5. **C**: Factor Structure	77092.403	29289	1	5.422	4	.25	-2.58
6. **E**: Simplex Structure	77131.185	29291	1	44.204	6	.00	32.20
7. **E**: Factor Structure	77100.577	29289	1	13.596	4	.01	5.60
8. Simplified model **A**: Simplex Structure **C**: Factor Structure **E**: Saturated	77098.961	29295	1	11.980	10	.29	-8.02
9. Simplified model – No Sex Differences **A**: Simplex Structure **C**: Factor Structure **E**: Saturated	77130.102	19320	8	31.141	25	.18	-18.86

[a] c.t.m. = compared to model

Table 3.3. - continued

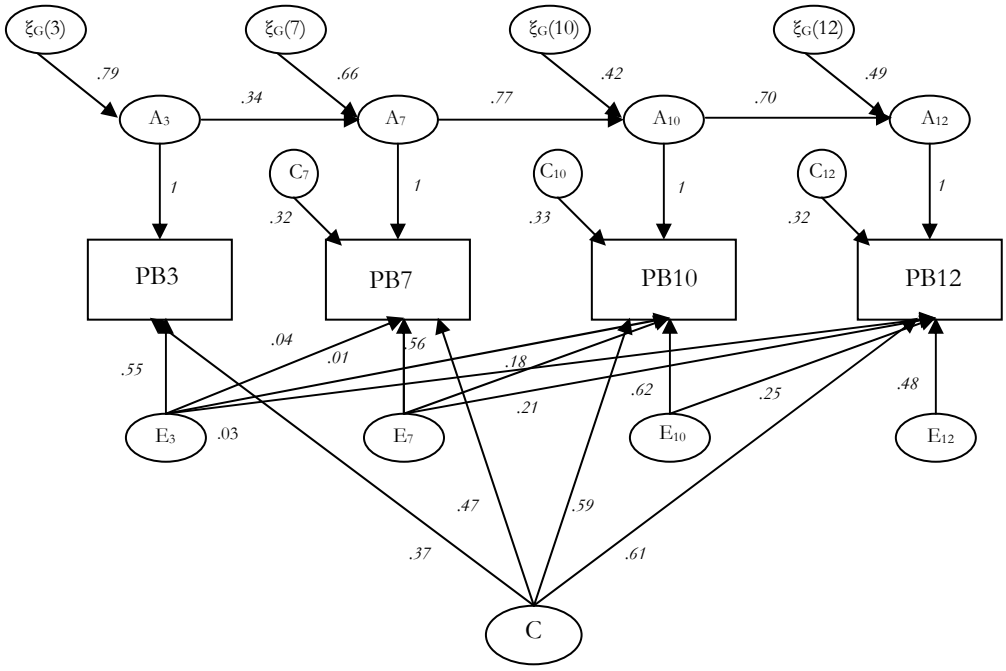
Multivariate model fitting for Externalizing problem behavior at age 3,7,10 and 12

EXTERNALIZING							
MODEL	**-2LL**	**df**	**c.t.m.**	**χ^2**	**df**	**p**	**AIC**
1. Saturated model **A**: Saturated **C**: Saturated **E**: Saturated	82410.846	29569					
2. **A**: Simplex Structure	82420.190	29575	1	9.344	6	.16	-2.66
3. **A**: Factor Structure	82424.485	29573	1	13.639	4	.00	5.64
4. **C**: Simplex Structure	82427.391	29575	1	16.545	6	.01	.455
5. **C**: Factor Structure	82415.336	29573	1	4.49	4	.34	-.351
6. **E**: Simplex Structure	82455.500	29575	1	34.654	6	.00	22.65
7. **E**: Factor Structure	82433.770	29573	1	22.924	4	.00	14.92
8. Simplified model **A**: Simplex Structure **C**: Factor Structure **E**: Saturated	82429.340	29579	1	18.494	10	.05	-1.51
9. Simplified model – No Sex Differences **A**: Simplex Structure **C**: Factor Structure **E**: Saturated	82506.974	29604	8	77.634	25	.00	27.63

For Externalizing problem behavior (Table 3.3) a model with a simplex structure for additive genetic influences and the saturated model for shared en nonshared environmental influences (model 2) did not gave a significantly worse fit than the full saturated model (model 1) (χ^2=9.344, Δdf = 6, p=.16). Further, shared environmental influences display a factor structure (model 5). For nonshared environmental influences on both Internalizing and Externalizing problem behavior neither the simplex nor factor model gave a satisfactory fit. The implication is that the processes that account for the nonshared environmental contribution were complex and could not be captured by one of our relative simple models. For Externalizing problem behavior sex-differences seem to be significant (model 8).

Figure 3.2.

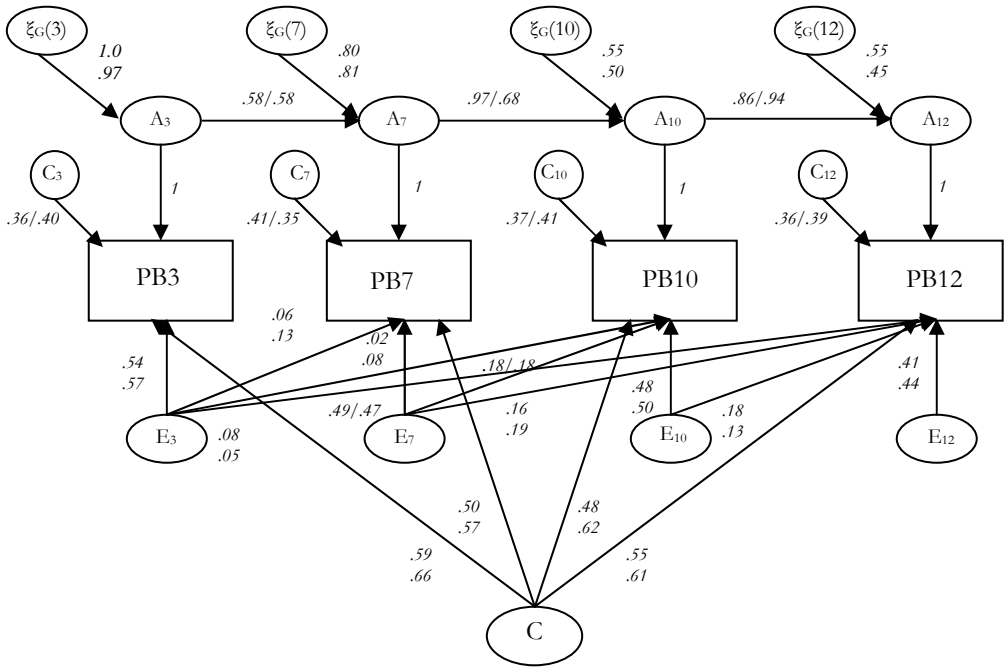
The best fitting model for Internalizing problem behavior, with a simplex structure for additive genetic factors, a common factor with age-specific influences for shared environmental factors, and a Cholesky Decomposition for nonshared environmental factors. (note: the variances of all latent factors are fixed to unity)



The parameter estimates based on the best fitting model for Internalizing and Externalizing problem behavior are represented in Figure 3.2 and 3.3. The percentages of the total age specific variance and the total between age covariances explained by additive genetic, shared environmental, and nonshared environmental factors based on the best fitting reduced models are presented in Table 3.4. In this table the total genetic variance at each specific age is decomposed in variance due to innovation and variance due to transmission from a previous age. The total shared environmental variance is decomposed into variance due to the common underlying factor and variance due to age-specific shared environmental influences. For nonshared environmental influences stability and change in variance specific to a certain age, and variance as a result of nonshared environmental factors from a previous age.

Figure 3.3.

The best fitting model for Externalizing problem behavior in boys (first number) and girls (second number), with a simplex structure for additive genetic factors, a common factor with age-specific influences for shared environmental factors, and a Cholesky Decomposition for nonshared environmental factors. (note: the variance of all latent factors are fixed to unity)



Comparable to the univariate model fitting results, both additive genetic and shared environmental factors are important in explaining individual differences in Internalizing and Externalizing problem behavior (Van der Valk *et al.*, 1998; 2001; 2002; Bartels *et al.*, 2002a; 2002d). A decrease in additive genetic effects and an increase in shared environmental effects is observed for Internalizing problem behaviors (e.g. A Int age 3: 59%; A Int age 12: 32%). These influences seem to stabilize from age 7 onwards. For Externalizing behavior less change in additive genetic and shared environmental influences is observed, although an increase in additive genetic influences for boys between age 3 and 10 can be seen in Table 3.4.

Important in Table 3.4 (off-diagonal) are the influences of additive genetic, shared environmental and nonshared environmental influences on the covariances. Stability, represented in these covariances, can be explained both by additive genetic and shared

environmental influences. Nonshared environmental influences seem of no significance for stability in problem behavior, represented by very low influences on the covariances.

Table 3.4.

Percentages of the total variances (diagonal) and covariances (off-diagonal) for Internalizing and Externalizing problem behavior explained by additive genetic, shared environmental and nonshared environmental components based on the best fitting models.

INTERNALIZING												
	A				**C**				**E**			
	3	**7**	**10**	**12**	**3**	**7**	**10**	**12**	**3**	**7**	**10**	**12**
3	.59				.00 .13				.28			
7	.52	.38[a] .06			.43	.09 .19			.05	.27 .01[c]		
10	.43	.51	.14 .22		.56	.36	.09 .25		.01	.13	.28 .03	
12	.32	.40	.38	.19 .18	.63	.42	.40	.08 .29	.04	.17	.22	.18 .08
EXTERNALIZING BOYS												
	A				**C**				**E**			
	3	**7**	**10**	**12**	**3**	**7**	**10**	**12**	**3**	**7**	**10**	**12**
3	.57				.07 .20				.16			
7	.64	.39 .20[a]			.33	.11 .15			.03	.14 .01[c]		
10	.66	.74	.16 .49		.33	.19	.07 .13		.01	.07	.13 .01	
12	.57	.69	.73	.16 .48	.38	.23	.18	.07 .16	.05	.07	.08	.09 .04
EXTERNALIZING GIRLS												
	A				**C**				**E**			
	3	**7**	**10**	**12**	**3**	**7**	**10**	**12**	**3**	**7**	**10**	**12**
3	.50				.08 .24				.18			
7	.54	.40 .19			.38	.07 .20			.07	.13 .01		
10	.45	.60	.16 .29		.50	.32	.11 .25		.06	.09	.16 .03	
12	.44	.58	.58	.13 .38	.52	.33	.33	.09 .24	.04	.09	.09	.12 .04

[a] the first number represents variance explained by age specific influences, the second number represents common influences specific to the underlying developmental mechanism

The only exception is the covariance between age 10 and 12 for Internalizing behavior, from which 22% is accounted for by nonshared environmental influences. On average, stability of Internalizing behavior over the years is for 47% ((52% + 43% + 32% +51% + 40% + 38%) /6) explained by shared environmental factors. 43% of this stability can be explained by additive genetic factors. The remaining 10% can be explained by nonshared environmental influences.

For Externalizing behavior in both boys and girls, though, additive genetic factors seems to be the main source of stability. For Externalizing behavior in boys 67% of the stability is explained by additive genetic factors. Twenty-seven percent and six percent of the stability is explained by shared and nonshared environmental influences, respectively. For Externalizing problem behavior in girls 53% of the stability over the years is explained by additive genetic factors. Shared environmental influences account for 40% of the stability, while 7% is explained by nonshared environmental influences.

DISCUSSION

To understand the development of Internalizing and Externalizing problem behavior structural equation modeling techniques were used in a large longitudinal sample of Dutch twins. This longitudinal study with large sample sizes and four measurement occasions gave us the unique opportunity to distinguish between a simplex and a common factor process underlying the development of Internalizing and Externalizing problem behavior. The simplex model assumes that successive levels of functioning were causally linked and that earlier experiences and/or genetic effects affected later maladjustment. The factor model related continuity in problem behavior to stable underlying environmental and/or genetic factors.

Our phenotypic correlation structure seems to be consistent with phenotypic stability coefficients reported in large scale longitudinal studies. Verhulst and Van der Ende (1992a, 1992b) studied stabilities in problem behavior in a sample of 936 Dutch 4- to 11-year old children. This study is especially interesting because children were almost the same age range and from the same Dutch population as the children in the present article. The average observed stability coefficients for the two-four-, and six-year time intervals were, respectively, .53, .48, .42. The stability coefficients in our study and this previous comparable study suggest a simplex pattern as the underlying developmental process. However, the stability between the distinct ages is higher than would be expected based on a simplex structure solely. This same pattern of stability coefficient is also found in a national sample of 16,000 British children at ages 7, 11, and 16 years (Ghodsian *et al.*, 1980). In comparison with the present study, though, children were somewhat older, and the interval covered the onset of puberty and transition to high school. The previous

studies could not distinguish between genetic and environmental influences on the developmental process. These different sources of variances may display a distinct developmental pattern. An important feature of the present longitudinal twin studies is the possibility to investigate the developmental pattern of each source of variance independently.

Stability in the development of Internalizing problem behavior can be explained by additive genetic factors accounting for 43% of the stability on average. Another factor explaining stability in the development of Internalizing problem behavior is an underlying common factor for shared environmental influences accounting for 47% of the total stability over the years. 10% of the stability of Internalizing problem behavior over the years is accounted for by nonshared environmental factors. Change is mainly accounted for by nonshared environmental influences. Genetic innovation factors and small but significant age specific influences of shared environment account for some change as well.

A comparable pattern for stability is found for Externalizing problem behavior. Stability is represented by additive genetic transmission factor explaining 67% of the stability over the years, on average, for boys and 53% of the stability over the years, on average, for girls. Stability is further accounted for by a common shared environmental factor explaining 27% and 40% of the total stability for boys and girls, respectively. Change in Externalizing behavior in both boys and girls can be mainly explained by nonshared environmental influences. Genetic innovations and age specific shared environmental influence account for some change in Externalizing problem behavior over the years.

The finding of a simplex pattern for additive genetic influences on Internalizing and Externalizing behavior seems remarkable. However, for both Internalizing and Externalizing behavior the phenotypic cross correlations are lower for longer intervals. This structure suggests a simplex pattern. This simplex structure, though, cannot be the sole mechanism because the stability between the distinct ages is higher than would be expected based on a simplex structure. The finding of a simplex structure for additive genetic influences in this study explained the finding of the developmental pattern in the phenotypic correlation. This simplex pattern for additive genetic factor accounts for the lower cross correlations for longer intervals. The higher than expected stability between the distinct ages based on a simplex structure solely, though, can be explained by the common underlying factor for shared environmental influences.

Further, several authors have pointed out that although all genes are present from conception onwards, this does not necessarily imply that genetically influenced traits are stable over time (Plomin, 1986). This is because not all genes are important all the time, and effects of specific subsets of genes may be age dependent. This fluctuation of gene

activity could also explain the relatively large innovation at age 7 (38% of total variance), and the lower to stable innovation at age 10 (14%) and 12 (19%). Between age 3 and 7 biological changes are expected due to development. Between age 7 and 12, however, no major biological changes are expected, represented in low but significant genetic innovation. Further, the period between 3 and 7 contains children's transition to school. During this transition children must cope with many new demands like meeting academic challenges, learning school and teacher expectations, adjusting to the daily routine of a school class (Barth and Parke, 1993; Cowan *et al.*, 1994; Ladd and Price, 1987). In a longitudinal study on the development of cognitive abilities, Fulker and colleagues (1993) suggested that genetic innovation at age 7 may be due to novel environmental challenge of schooling. Continuing the longitudinal study would be informative on the expected genetic changes due to puberty.

The idea of transmission is quite common and present in many developmental theories such as the psychoanalytic or attachment theory (Lamb and Nash, 1989). In addition, developmental concepts like *critical periods* or *developmental tasks* also refer to a process in which outcomes of certain phases affect future function. No previous study did find this transmission to be accounted for by genetic factors rather than environmental influences.

The finding for shared environmental influences on Internalizing and Externalizing problem behavior is less surprising. Our study indicates that, besides a continuing influence of shared environmental factors, age specific influences are present. These age-specific effects were significant but the proportion of variance explained is much smaller compared to the proportion explained by the shared environmental factor common to all ages. This common factor could be accounted for by stable familial factors such as SES, as this important shared environmental aspect is not sensitive to large changes over a time-span of 9 years. Aspects outside the family environment, like friends or being a member of a sports club or school, might also cause similarities between two children of a twin pair during childhood, but could be age specific rather than continuous throughout development. An explanation for age-specific shared environmental at ages 7, 10 and 12 could be the change of teacher at every level in Dutch elementary schools. Information on 'same' or 'different' teacher for both children of a large sample of 7, 10, and 12 year old twin pairs indicate that in 63% of the cases, on average, both children of a twin pair are taught by the same teacher, whereas 37% go to separate classes. This ratio makes teacher or classroom environment a shared environmental influence for the majority of the children. Since, in the Dutch school system children move to a different teacher each school year, this results in a lack of continuity in this particular aspect of shared environment. So, these shared but age-specific experiences within the classroom may be represented by the age-specific factors as specified significant in the best fitting model.

The finding in the present study that shared environmental influences are represented by a common factor and time specific influences replicates the results of developmental studies in other areas (Bartels *et al.*, 2002b). It indicates that there could be a very stable set of shared environmental influences that causes problems to persist over the years. Similar results seem to be suggested by a number of epidemiological studies showing that problems tend to continue in families with ongoing family adversities like marital stress, negative maternal control, and maternal depression (Campbell *et al.*, 1991; Egeland *et al.*, 1990; Richman *et al.*, 1982). Thus, not family adversity as such but its persistence predicts chronic problems (Campbell, 1994). It should also be mentioned that although ongoing family adversity may indeed represent shared environmental influences, parts of it relation with continuity in problem behavior might be explained by genetic influences. This could be due to genetic factors that are shared by parents and children and influence both the family environment and children's behavior (Braungart-Rieker *et al.*, 1995; Plomin, 1995; Plomin et al, 1994; Rowe, 1981, 1983).

Another explanation for the significance of shared environmental influences could be rater bias. Sources of rater bias are stereotyping, employing different normative standards, or having certain response styles, i.e. judging problem behaviors more or less severely. Is expected that rater bias in this sense will be a continuous process influences the ratings at all ages. Less obvious, but not erasable is the fact these types of bias may change over time, for instance mothers change their opinion on certain kinds of behavior leading to change in rating style. This change of rating style could show up as age specific shared environmental influences at the distinct ages. In order to solve this uncertainty about continuous or age specific influences of shared environmental influences and to distinguish 'real' shared environmental influences form rater bias, longitudinal psychometric models, making use of mother and father ratings, should be used. These models assume that, in addition to assessing similar aspects of the child's behavior, each parent assesses an unique aspect of their child's behavior. This results in aspects of the child's behavior similar assessed by both parents, representing 'real' unbiased behavior, besides factors of rater bias and unique views of each parent. If the shared environmental influences are a result of maternal rating style, it should show up in the unique shared environment influences instead of shared environmental influences on the part of behavior similar assessed by both parents.

Nonshared environmental influences were substantial at each age, and contributed mainly to change in children's problem behavior. For the covariance between ages 7, 10, and 12 in Internalizing behavior, however, nonshared environmental factors seem to be of significant influence. Possible examples of nonshared environmental influences include illness, trauma, fluctuations in mood and state, and peer group influences (Plomin and

Daniels, 1987; Rowe *et al.*, 1994). Findings from this study imply that these adverse experiences are important and that they are mostly of transient nature and children 'recover' from them, but sometimes also exert long-lasting effects.

Limitations of the study and clinical implications

Certain aspects in studying the etiology of childhood psychopathology using twin pairs and parental ratings need to be considered. A previous cross-sectional study on contrast effects for Attention Problems in a comparable sample of Dutch twins detected a rater contrast effect at age 3. The authors hypothesized that the contrast effect represented a maternal rater bias effect that is dependent on the age of the twins (Rietveld *et al.*, 2002). Further, a study in an overlapping sample of 3-year-old Dutch twin pairs showed evidence of sibling interaction for Externalizing behavior. The interaction proved to be in a cooperative manner, with twins reinforcing each other's behavior (Van der Valk *et al.*, 1998b)

In general contrast effects (b) may be considered as a social interaction between siblings (Carey, 1986; Eaves, 1976) or an effect introduced by the rater (Neale and Stevenson, 1989). In the former case, the behavior of one twin has a certain effect on the behavior of his or her co-twin. This effect can be either cooperative or competitive. In the latter case, when parents are asked to evaluate and report upon the problem behavior of their children, they may very well compare the twins' behavior against one another, despite instructions on the questionnaire form. In this way, one twin becomes some kind of standard by which the behavior of the co-twin is rated. A significant contrast effect is implied when MZ variances and DZ variances are heterogeneous. In addition to heterogeneity of MZ and DZ variances, the presence of a contrast effect leads to a pattern of MZ and DZ correlations that is inconsistent with additive genetic sources of variances.

In both Table 3.1 (variances) and Table 3.2 (twin correlations) no indication for a contrast effect is observed. However, to get insight in the influences of contrast effects on Internalizing and Externalizing problem behavior, we fitted cross-sectional models taking this interaction parameter into account for Internalizing and Externalizing problem behavior at ages 3, 7, 10, and 12. For Internalizing behavior in girls at age 7 (b= .07) and for Externalizing behavior in both girls and boys at age 3 (b=.11), solely, significant influences of rater contrast are found. Because of the fact that these are the only significant finding, contrast effects are not considered in the longitudinal model fitting procedures.

Further, our analyses were performed on a non-clinical sample. Assuming that psychopathology is caused by environmental hazards or pathogenic genes that are qualitatively distinct from those that cause variation in the normal range (Rutter *et al.*,

1990), our result would have little clinical importance. There is, however, evidence that clearly suggests links between normal and abnormal behavior. First, several CBCL studies have shown correlations between behavior problem syndromes and DSM diagnoses (Costello *et al.*, 1985; Edelbrock and Costello, 1988; Ferdinand *et al.*, 1999; Kasius *et al.*, 1997). This convergence indicates that behavior problem syndromes as studied in this article must be relevant for psychiatric conditions. Second, several studies supported the view that the sources of normal variation may also affect psychopathology in children and adolescents. So latent class analyses have been used to identify subgroups of individuals with normal or pathological behavior (Eaves *et al.*, 1993; Hudziak *et al.*, 1998; Neuman *et al.*, 1999). Results tend to suggest that these groups differ in degree rather than in kind. Furthermore, using methods from item response theory, Van den Oord and colleagues (2002) found that liability distributions for behavior and emotional problems show very little or no evidence of non-normality. This also seems to suggest that psychopathology may often be an extreme on the same continuum that describes variation in the normal range. Thus, although we used a non-clinical sample, it can be argued that our longitudinal analyses are also important for understanding psychopathology.

Longitudinal behavior genetic analyses provide knowledge about the mechanisms underlying stability and change in problem behavior. Our finding of different developmental patterns for the distinct sources of variance (A, C, and E) has important implication for the prevention of later maladjustment. The shared environmental influences, for instance, exert a continuous influence from their time of onset. So, the children who continue to experience adverse shared environment are at risk for later maladjustment. For additive genetic influences, parts of previous effects are transmitted to later ages. However, genetic influence is less static due to new genetic influences that come into play at each age. Nonshared environmental influences seem to be important for age-specific behavior problems and have almost no developmental significance. This implies that influences of nonshared environment are important but that they are mostly of transient nature and specific to a specific moment in time.

Further, multivariate behavior genetic analyses of pattern of problem behavior make it possible to distinguish disordered children into groups that have mainly a genetic, shared environmental, or a nonshared environmental etiology to make the crucial differential diagnosis (e.g. Boomsma *et al.*, 1990; Van den Oord *et al.*, 2000). In combination, with the knowledge about mechanisms that underlie stability and change, such a subdivision might be useful for prevention. For instance, for both Internalizing and Externalizing problem behavior continuing genetic and shared environmental effects were import for stability. When these results are generalized it implies that especially children with high genetic liability or children who continue to experience adverse shared environment are at risk for

later maladjustment. For these children, a 'wait and see' policy would be inappropriate and an active intervention would be required.

Appendix 3.1

Twin correlations and cross correlations for Internalizing (above diagonal) and Externalizing (below diagonal) problem behavior for the five zygosity groups.

MZM	**3**	**7**	**10**	**12**
3	.712[a] (859)\ .840[c] (859)	.283 (694)	.266 (408)	.240 (227)
7	.498 (701)[b]	.693 (853)\ .847 (865)	.529 (432)	.420 (237)
10	.463 (417)	.665 (447)	.659 (498)\.857 (514)	.487 (253)
12	.461 (228)	.616 (239)	.695 (254)	.734 (273)\.875 (276)
DZM	**3**	**7**	**10**	**12**
3	.408 (868)\ .567 (867)	.250 (684)	.271 (397)	.248 (.192)
7	.327 (697)	.460 (799)\.544 (822)	.370 (396)	.239 (182)
10	.287 (397)	.400 (404)	.481 (465)\ .535 (467)	.405 (195)
12	.266 (198)	.314 (190)	.399 (202)	.451 (217)\ .531 (405)
MZF	**3**	**7**	**10**	**12**
3	.734 (1006)\ .832 (1003)	.399 (821)	.316 (502)	.397 (250)
7	.541 (829)	.734 (975)\ .861 (989)	.513 (521)	.473 (260)
10	.437 (509)	.634 (531)	.680 (602)\ .811 (609)	.540 (276)
12	.410 (257)	.600 (270)	.612 (283)	.719 (301)\.839 (310)
DZF	**3**	**7**	**10**	**12**
3	.428 (810)\ .564 (813)	.275 (646)	.284 (359)	.230 (176)
7	.369 (663)	.533 (778)\.553 (802)	.387 (383)	.354 (188)
10	.337 (364)	.396 (395)	.507 (442)\ .582 (450)	.329 (197)
12	.250 (180)	.286 (196)	.389 (202)	.535 (214)\ .553 (223)
DOS	**3**	**7**	**10**	**12**
3	.388 (1752)\ .542 (1752)	.243 (1324)	.286 (745)	.263 (341)
7	.340 (1342)	.509 (1550)\ .562 (1591)	.351 (718)	.342 (324)
10	.301 (748)	.377 (740)	.532 (871)\ .502 (883)	.384 (345)
12	.273 (352)	.408 (344)	.454 (361)	.592 (398)\ .605 (413)

[a] internalizing behavior, [b] number of complete twin pairs, [c] externalizing behavior

4

Disentangling genetic, environmental, and rater effects on Internalizing and Externalizing problem behavior in 10-year-old twins

---

This chapter is submitted as:
Bartels, M., Boomsma, D.I., Rietveld, M.J.H., Van Beijsterveldt, C.E.M., Hudziak, J.J., and Van den Oord, E.J.C.G. (2002). Disentangling Genetic, Environmental, and Rater Effects on Internalizing and Externalizing Problem Behavior in 10-year-old Twins. *Psychological Medicine*, submitted.

ABSTRACT

Previous studies have emphasized the importance of rater issues in studying the etiology of variation in Internalizing and Externalizing problems in children. Earlier results indicate only moderate agreement between parents, and assume that parents assess an unique aspect of their child's behavior. In comparable samples of younger children, additive genetic effects are the main factor explaining individual differences in both Internalizing and Externalizing behavior. It is unknown whether this pattern of rater influences and variance decomposition will be consistent in older children. Child Behavior Checklists (Achenbach, 1992), filled in by mothers and fathers, were collected in a sample of 2,956 Dutch 10-year-old twin pairs. The etiology of individual differences in Internalizing and Externalizing syndromes was examined using a model that corrected for possible rater bias, unique rater effects and unreliability. The best fitting model suggested that disagreement between the parents is not merely the result of unreliability and/or rater bias, but each parent also provides unique information from his/her own perspective on the child's behavior. Significant influences of additive genetic, shared environmental and nonshared environmental factors were found for Internalizing and Externalizing syndromes. Besides parental agreement, unique parental views on their children's behaviors seem to be significant at age 10. These results are in line with the findings in comparable samples of Dutch twins at ages 3 and 7 years. Additive genetic factors remain important as a source of individual differences in Internalizing and Externalizing problem behavior. Shared environmental influences are also not to be minimized. The changes in variance decomposition that occurs, occurs mainly between age 3 and 7. No major changes are observed between age 7 and age 10.

INTRODUCTION

Parental descriptions are often used to collect information about a child's behavioral and emotional problems. A meta-analysis by Achenbach, McConaughy, and Howell (1987) showed a mean correlation of .60 between maternal and paternal ratings of the same child. The high interparent correlation shows that parents can provide meaningful information about their child's behavior, for if parental ratings would reflect nothing but error the correlations between their ratings would be close to zero. On the other hand this interparent correlation is less than perfect. This may be explained by different forms of rater bias (a tendency of a rater to over-or underestimate behavioral problem scores consistently compared to the mean of all raters, e.g. a result of different normative standards or response tendencies) and unreliability. Sources of rater bias are stereotyping, employing different normative standards, or having certain response styles, i.e. judging problem behaviors more or less severely. Because these types of bias may differ between raters, they may lead to disagreement between raters. Unreliability can become an important source of disagreement when raters cannot give an accurate description about relevant behaviors. For instance, evidence is found that parents may be relatively insensitive to affective disturbances in children (Angold et al., 1987). Another possibility is that parents are not assessing exactly the same behavior in their children. It is know that different raters can provide, each from their own perspective, somewhat different but valid and complementary information about the child's functioning (Achenbach *et al.*, 1987). Loeber *et al.*, (1989), for instance, found that children's reports on their conduct problems tended to complement the information provided by adults.

It is difficult to draw conclusions about the processes underlying the (dis)agreement between parental ratings on the basis of the parental intercorrelations alone. Genetically informative data are helpful in this respect allowing, due to their special properties, the evaluation of different hypotheses about the (dis)agreement in parental ratings. Models can be fitted to the data to test whether parental disagreement is caused by unreliability and rater bias, or involves the fact that parents provide unique information about their children's behavior. A correct representation is not only important from a substantive point of view, but also to obtain more accurate estimates of genetic and environmental effects. For instance, rater bias will cause shared environmental effects to be overestimated, measurement error will magnify the estimate of nonshared environment. The use of multiple raters makes it possible to disentangle these rater effects from variance caused by the child's behavior so that parameter estimates are less biased and have a clearer interpretation.

To study agreement and disagreement between parental ratings, Hewitt *et al.* (1992) proposed so-called Rater Bias and Psychometric models that combine data of two raters and can be estimated using genetically informative data. The Rater Bias model assumes that parents

assess the same behaviors in the child and have a common understanding of the behavioral descriptions. This may apply when both parents are equally confronted with the behaviors shown by the child (for instance at home). Disagreement between the raters is regarded as error, resulting from rater bias and/or unreliability. In addition to assessing similar aspects of the child's behavior, the Psychometric model assumes that each parent assesses an unique aspect of the child's behavior. This will occur when the parent observes the child in distinct situations or is exposed to distinct samples of the child's behavior. For instance, the parent who usually brings the child to school may be more familiar with the child's behavior outside the home. Moreover, each parent may interact differently with the child (Achenbach et al., 1987). These unique interactions between a parent and a child may allow each parent to provide additional information about the child's behavior, apart from the information on which they both agree. Disagreement in this model does not merely arise from unreliability and/or rater bias, but also because each parent contributes, from his own perspective, different but valid information on the child's functioning. The psychometric model tests this possibility by examining whether there are significant genetic effects on the unique part of each parent's rating. If the behaviors uniquely rated by the parents are shown to be influenced by the genotype of the child, the parent must have been assessing a 'real' but unique aspect of the child's behavior.

A number of quantitative genetic studies have used the Child Behavior Checklist (CBCL) (Achenbach, 1991, 1992) to examine genetic and environmental effects on children's problem behaviors (Silberg *et al.*, 1994; Edelbrock *et al.*,1995; Schmitz *et al.*, 1995; Van den Oord *et al.*, 1996; Zahn-Waxler *et al.*, 1996; Gjone and Stevenson, 1997; Leve *et al.*, 1998; Van der Valk *et al.*, 1998a; Van der Valk *et al.*, 1998b; Hudziak *et al.*, 2000). Yet, only a few studies employed models that incorporated rater differences. Rowe and Kandel (1997) administered the CBCL to mothers and fathers for their oldest two offspring (aged 9 to 17) in 76 families. The subjects, however, were nontwin siblings rather than twins. Hence, estimation of separate genetic and environmental components of trait variance was impossible. The combination, though, of three informants (mother, father, self-report) and the rating of two children per family, allowed to disentangle rater effects from variance caused by a common understanding of the behavioral description in parents. Their models demonstrated that mother and father ratings contained a substantial individual view component, but parents also assessed similar aspects of the child's behavior. Hewitt and colleagues (1992) fitted Rater Bias and Psychometric models to parental ratings of the Internalizing scale (CBCL) for 983 twin pairs. They found that both for their prepubertal cohort (8 to 11 years) and for their pubertal cohort (12 to 16 years) the Psychometric model fitted the data better than the Rater Bias model. Van der Valk and colleagues (2001, van der Valk *et al.*, in Press) also found that the Psychometric model fitted their data significantly better than the Rater Bias model at both ages 3- and 7.

Thus these studies indicated that disagreement between parental ratings is partly caused by mothers and fathers assessing different aspects of the child's behavior.

In the present study we fitted Rater Bias and Psychometric models to data for the Internalizing and Externalizing scale of the CBCL. The sample consisted of 2,956 Dutch 10-year-old twin pairs. The first aim of this study is to fit Rater Bias and Psychometric models. Results of previous studies in comparable samples of Dutch twins indicated a Psychometric model as best fitting model. Achenbach *et al.* (1987) observed, however, that the correlation between similar informants (e.g. parents) decreased with age of the child. On explanation is a decrease in the quality of parent ratings. Parents mainly interact with their children in the home environment. However, as children become older other social context such as school and the peer group become relatively more important. Consequently, it becomes more difficult for parents to assess problem behaviors in their children. Such a possible decrease could result in a better fit of our Rater Bias model compared to the Psychometric model. Another explanation for lower parental agreement in older children is that parent-child relations become more individual and specialized over the years. For instance, the roles of mother and father may become more differentiated and they may engage in different activities with their children. Such a change would imply that the unique view of the parent increases. This specialization would suggest, in line with the findings at age 3 and 7, a Psychometric model at age 10.

A second aim is to use the best fitting model to estimate influences of genetic and environmental components on Internalizing and Externalizing Problem Behavior at age 10. Comparison of the results of this study to the results of comparable studies in 3-and 7-year-old Dutch twins gives the opportunity to disentangle 'real' behavioral development from changes in rater effects. The large sample of twin pairs used provided the power necessary to be able to detect possible small changes.

METHOD

Subjects

All participants were registered by the Netherlands Twin Registry (NTR), kept by the Department of Biological Psychology at the Vrije Universiteit in Amsterdam. Of all multiple births in the Netherlands, 40-50% is registered by the NTR (Boomsma *et al.*,1992; Boomsma, 1998; Boomsma *et al.*, 2002). For this study, data from twins from the birth cohorts 1986 - 1991 were used. Questionnaires were mailed to families within three months of the twins' tenth birthday. After two to three months reminders were sent and four months after the initial mailing persistent non-responders were contacted by phone. Families whose addresses were not available were included in the nonresponse group. 140 twin pairs were excluded because either one or both of the children had a disease or handicap that interfered severely

with daily functioning at age 10 or at a younger age. Finally, the analyzed sample consists of 2,956 mother ratings and 2,234 father ratings.

Zygosity was determined for 620 same-sex twin pairs by DNA or blood group polymorphisms. For all other same-sex twin pairs zygosity was determined by discriminant analysis, using questionnaire items. Parents were asked how much the twins resembled each other in facial structure, hair color, facial color, eye color, and whether they were ever mistaken for each other by the parents themselves, by family, or by strangers. They were also asked if the twins were as much alike as two peas in a pod, whether it was difficult for the parents to separate the twins on a recent picture, and how similar the twins' hair structure was. (For details see Rietveld *et al.*, 2000).

This left a sample of 519 monozygotic males (MZM), 471 dizygotic males (DZM), 618 monozygotic females (MZF), 458 dizygotic females (DZF), and 890 dizygotic opposite sex (DOS) twin pairs. In general, mothers' response rate outnumbers fathers' response rate. Therefore the data could be further divided into twin pairs for which both mother and father had replied (400 MZM, 347 DZM, 470 MZF, 348 DZF, and 669 DOS) and twin pairs for which only mothers had replied (119MZM, 124 DZM, 148 MZF, 110 DZF, and 221 DOS). Because of a relative small amount of families from which only fathers replied (N=28) these families were not used in the analyses.

Measures

The Child Behavior Checklist (CBCL 4-18) (Achenbach, 1992) is developed for parents to score the behavioral and emotional problems of their 4-to-18-year-old children. It consists of 120 problem items that are scored by the parents on a 3-point scale based on the occurrence of the behavior during the preceding 6 months: 0 if the problem item was not true, 1 if the item was somewhat or sometimes true, and 2 if it was very true or often true. The syndrome scales were composed according to the 1991 profile (Achenbach, 1991). Dutch syndrome scales and comparability with the syndrome scales as developed by Achenbach are reported in Verhulst *et al.*, (1996). In this manual the two broadband scales Internalizing (Int) and Externalizing (Ext) are analyzed. The Internalizing scale consists of the Anxious/Depressed, Somatic Complaints and Withdrawn subscales. The Externalizing scale consists of the Aggressive and Rule Breaking Behavior subscales. For the Internalizing scale subjects were only included if not more than 3 items were missing for the Anxious/Depressed scale, not more than 2 items were missing for Somatic Complaints and Withdrawn scales. For the Externalizing scale the inclusion criterion was not more than 3 items missing for the Aggressive and Rule Breaking Behavior scales. This ensured that the two syndrome scales were always composed of all problem behaviors loading on that scale.

The data were square root transformed to approximate normal distributions that are

required for maximum likelihood estimation. After transformation, all skewness and kurtosis indices were between -1.0 and 1.0, implying that not much distortion is to be expected (Muthén and Kaplan, 1985).

The twin method

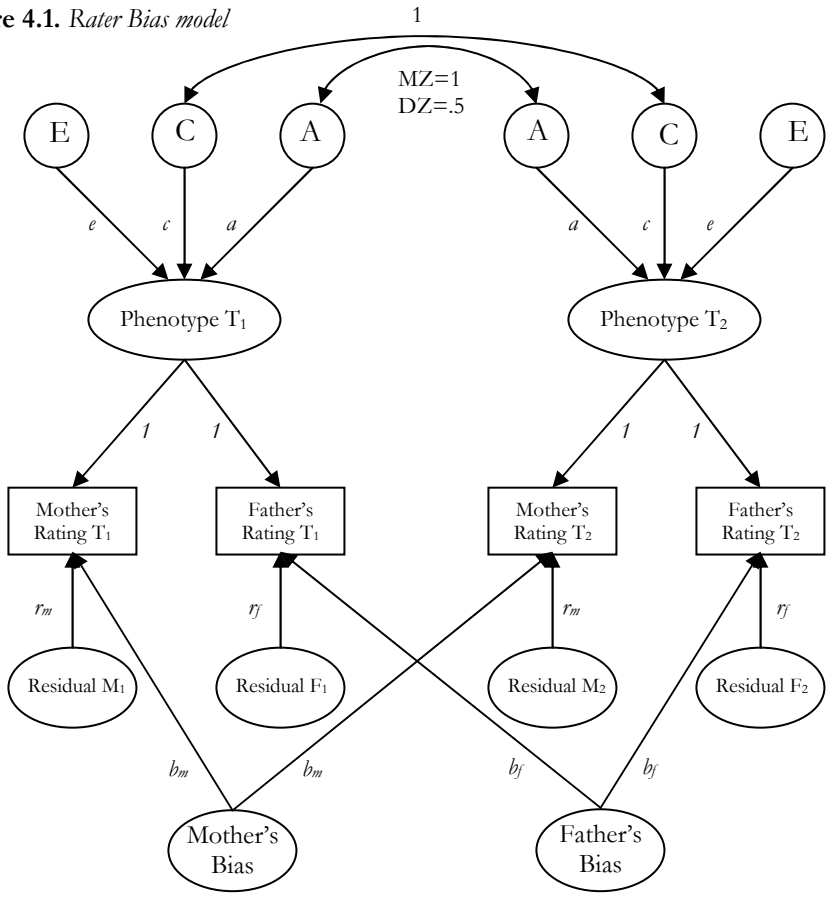
Data from monozygotic and dizygotic twins were used to decompose the variance in scores on the Internalizing and Externalizing scales into a contribution of the additive effects of many genes, environmental influences that are shared by twins (like style of parenting, socioeconomic level, or religion) and environmental influences that are not shared by twins (such as an illness, relationships with peers, or measurement errors). For a summary of the twin method, the various assumptions, and the plausibility of these assumptions see Martin and Eaves (1977); Eaves (1982); Kendler and Eaves (1986); Neale and Cardon (1992). An estimate of the additive genetic, shared environmental, and nonshared environmental influences can be derived from the resemblance between MZ twins who are genetically identical and DZ twins who share on average half of their genes. Genetic effects are indicated when the MZ twin correlation (r_{mz}) is higher than the DZ twin correlation (r_{dz}). Shared environmental effects are indicated if the twin correlations are larger than zero after the genetic effects are partialled out, and nonshared environmental effects are indicated if the correlation between MZ twins is smaller than 1. Assuming that the effects of genes are so that the genotypic correlation is .5 for DZ twins, the proportion of variance explained by each component can be calculated as follows: genetic variance = $2 \times (r_{mz} - r_{dz})$, shared environmental variance = $2 \times r_{dz} - r_{mz}$, and nonshared environmental variance = $1 - r_{mz}$. This approach, however, does not take into account sex differences and cannot easily be generalized to multivariate data. This approach also does not provide a goodness of fit of a certain model to the data.

To decompose the variance shared by both parents, the correlation between the twins rated by different raters (cross-correlation; e.g. Vader-Int in Twin 1 with Mother-Int in Twin 2) has to be used. This way, the variance is decomposed into additive genetic, shared environmental, and nonshared environmental contributions for which both parents agree. The decomposition can again be made by comparing the resemblance of MZ twins versus DZ twins. Genetic effects are indicated when the cross-correlation is higher for MZ twins compared to DZ twins. Shared environmental effects are indicated if the cross-correlations are larger than zero after the genetic effects have been partialled out, and a nonshared environmental contribution is indicated when the cross-correlations for MZ twins is smaller than the interparent correlation. Similar formulas to the ones discussed above for the variances can again be used to compute the contributions of each component: genetic contribution = $2 \times (r_{mz\text{-}cross} - r_{dz\text{-}cross})$, shared

environmental contribution = $2 \times r_{dz\text{-}cross} - r_{mz\text{-}cross}$, and nonshared environmental contribution = interparent correlation - $r_{mz\text{-}cross}$.

The above discussed formulas indicate that the whole variance-covariance matrix can be decomposed into a matrix of genetic variances and covariances, a matrix of shared environmental variances and covariances, and a matrix of nonshared environmental variances and covariances. Instead of decomposing each variance and covariance separately, it is preferable to make such a decomposition by fitting multivariate genetic models. For this purpose Hewitt *et al.* (1992) proposed a Rater Bias and Psychometric model.

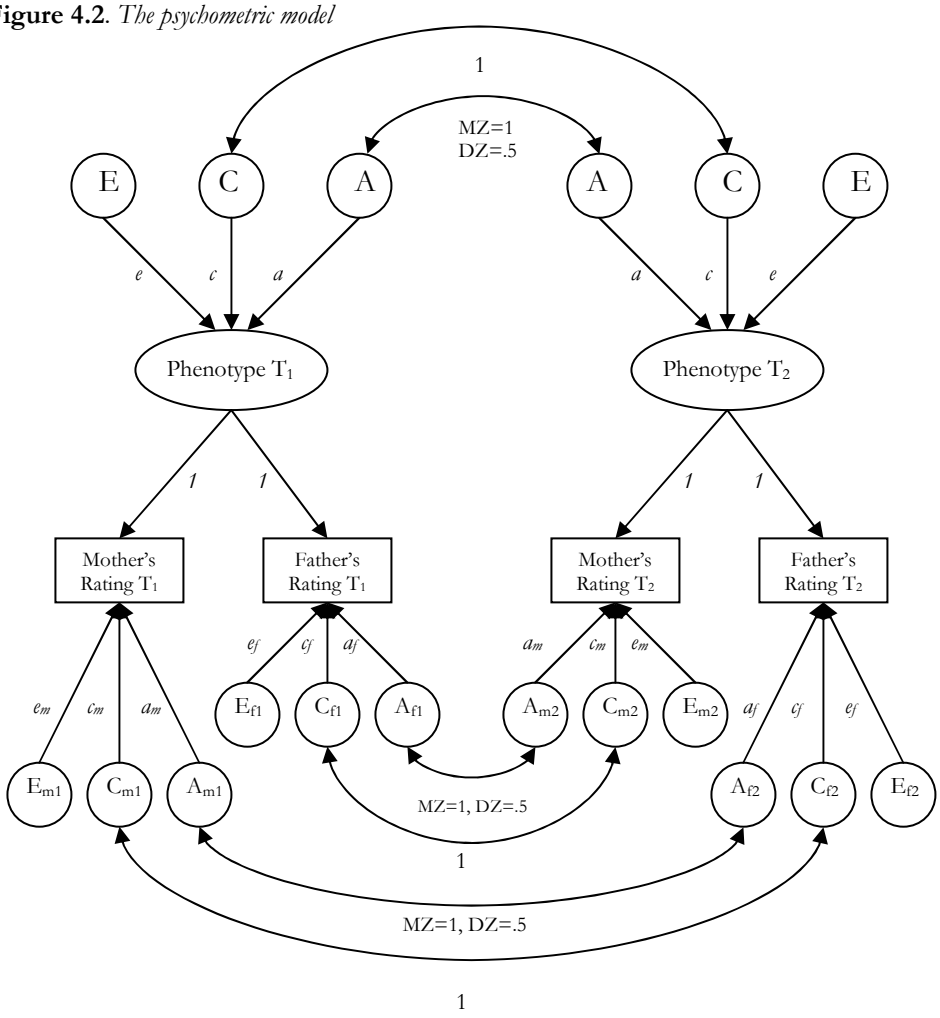
Figure 4.1. *Rater Bias model*



In the Rater Bias model (Hewitt *et al.,* 1992) (Figure 4.1) the phenotypes of the twins are a function of three common factors underlying the ratings of both mothers and fathers: a genetic factor (A), a shared environmental factor (C), and a nonshared environmental factor (E). In addition to these three common factors unique factors are modeled: a maternal rater bias factor, a paternal rater bias factor, and residual (unreliability) factors affecting each rating.

The influence of the common factors is assumed to be independent of the maternal and paternal rater bias and unreliability factors.

Figure 4.2. *The psychometric model*



The Psychometric model (Hewitt *et al.*, 1992) (Figure 4.2) also estimates the influence of a genetic (A), a shared environmental (C), and a nonshared environmental factor (E) common to the phenotypes of the twins as rated by both parents. In addition three unique factors, a genetic ($A_{m/f}$), shared environmental ($C_{m/f}$), and nonshared environmental factor ($E_{m/f}$) are estimated for the ratings of mother/father. Disagreement between parents in this model can be caused by unique behavioral views, leading to different but valid information of each rater.

These unique behavioral views can have their own unique influences, estimated in the unique additive genetic, shared environmental, and nonshared environmental factors. Disagreements can also be caused by rater bias, which will confound the unique shared environmental effects, or by unreliability, which will confound the unique nonshared environmental effects. The three common factors loading on the twins' phenotypes contain only reliable variance, causing the common nonshared environmental factor to contain only pure independent environmental effects (McArdle and Goldsmith, 1990) and the common shared environmental factor to contain only pure shared environmental effects.

Model fitting

The program Mx (Neale *et al.*, 1999) was used to analyze the data through a simultaneous analysis of the 4 × 4 variance-covariance matrices in the five zygosity by sex twin groups (MZM, DZM, MZF, DZF, DOS) where both mother and father ratings were available, and the 2 × 2 variance-covariance matrices in the five zygosity by sex twin groups with only mother ratings. The model describes the observed variance-covariance matrices adequately when the residual variance-covariance matrices are trivially small. A good model is indicated by a low non-significant χ^2 test statistic ($P > .05$). Apart from the χ^2 test statistic, Akaike's Information Criterion (AIC = χ^2 - 2 × degrees of freedom) was computed. The lower the AIC the better the fit of the model to the observed data.

Fitting the Rater Bias and Psychometric model of Hewitt *et al.* (1992) to the data showed which model described the processes involved in either agreement or disagreement between the parental ratings best. Monozygotic twin covariances and dizygotic twin covariances were modeled, assuming a correlation between the twins' shared environmental factors of 1.0, regardless of twin type, and a genotypic correlation of 1.0 for monozygotic twins and 0.5 for dizygotic twins. Estimates for male and female twins were allowed to differ. This model was further examined for possible simplifications. It was tested whether the common and/or unique factors could be removed from the model, whether estimates for boys and girls could be constrained to be the same, and if the unique factors for mothers and fathers could be constrained to be equal. The only factor that was never dropped from the model was the unique nonshared environmental factor, because measurement errors are estimated in this factor.

RESULTS

Description of the data

The untransformed mean problem scores and standard deviations of the twin sample and those of a Dutch community sample of 4-11-year-old children (Verhulst et al, 1996) are given in Table 4.1. For both the Internalizing and Externalizing scale, the ratings given to the twins

were quite similar to the ratings given to the Dutch community sample. Significance tests showed that boys did receive higher mother and father ratings than girls for the Externalizing scale. For this same scale, mothers gave higher ratings to their children than fathers did, implying possible rater differences. For the Internalizing scale mothers gave significantly higher ratings for girls and mothers gave higher ratings to their twin children than fathers did, implying possible rater differences for this scale as well.

Table 4.1.

Means (standard deviation) and sample size for the Internalizing and Externalizing scales in 10-year-old twins (per zygosity) compared to a 4-11 year- old Dutch community sample.

		INTERNALIZING			EXTERNALIZING		
		Mother	**Father**	**N (M/F)**	**Mother**	**Father**	**N (M/F)**
♂	**MZM**	4.59 (4.60)	3.50 (3.83)	1008/796	8.67 (7.43)	7.11 (6.65)	1029/798
	DZM	5.18 (5.44)	4.14 (4.57)	933/691	8.03 (7.34)	6.90 (6.31)	937/692
	DOS	4.66 (4.97)	3.56 (4.14)	877/665	7.69 (7.10)	6.51 (6.25)	885/667
	COM	4.52 (4.27)		579	8.26 (6.36)		579
♀	**MZF**	5.12 (5.00)	3.56 (3.95)	1216/943	5.91 (5.53)	4.81 (4.94)	1223/939
	DZF	5.35 (5.35)	4.30 (4.60)	893/688	5.90 (5.86)	5.04 (5.13)	905/696
	DOS	4.76 (4.91)	3.56 (3.70)	878/662	5.34 (5.37)	4.50 (4.79)	885/667
	COM	5.16 (5.02)		593	6.04 (5.57)		593

Note. MZM/DZM = Monozygotic/Dizygotic males, MZF/DZF = Monozygotic/Dizygotic females, DOS = Dizygotic opposite sex, COM= Dutch community sample, N (M/F) = number of children for Mothers (M) and Fathers (F).

The homogeneity of the variance was tested with Mx (Neale *et al.,* 1999). No differences could be found in the variances and covariances of MZM, DZM, MZF, DZF, and DOS, for the Internalizing scale. For the Externalizing scale MZM variance is equal to DZM variance and MZF variance is equal to DZF variance, however the variance for boys and girls, both MZ and DZ could not be set equal.

Twin correlations

Table 4.2 shows, for both the Internalizing and Externalizing scale, in the first and second column the correlations between the twins rated by the same rater (mother *or* father rated both children), and in the third and fourth column the cross-correlations between the twins each rated by a different rater (mother *and* father each rated one child). In the fifth and sixth column the interparent correlations between mothers and fathers are given, both for first and second

born twin. The interparent correlations were comparable for both first and second born twin for all zygosity by sex groups. On average, the interparent correlations for the Internalizing scale were .63, and for the Externalizing scale .73. This resembled the interparent correlations obtained in the Dutch norm group (Verhulst *et al.*, 1996).

The correlations between the first and second born twin both rated by mothers (M/M; first column) and those both rated by fathers (F/F; second column) can be used to obtain a first estimate of the genetic influences (A), the shared environmental influences (C), and the nonshared environmental influences (E) on the total variance.

Table 4.2.

Correlations (ratings given by the same rater), and cross-correlations (ratings given by different raters) between the twins and the interparent correlations, per zygosity, for 10-year-olds.

	INTERNALIZING						**EXTERNALIZING**					
	same rater		**different rater**				**same rater**		**different rater**			
	TWINS		**TWINS**		**INTER PARENT**		**TWINS**		**TWINS**		**INTER PARENT**	
	M/M	**F/F**	**M/F**	**F/M**	**O**	**Y**	**M/M**	**F/F**	**M/F**	**F/M**	**O**	**Y**
MZM	.66	.66	.40	.46	.64	.63	.86	.86	.67	.66	.77	.73
DZM	.48	.49	.26	.36	.67	.63	.54	.57	.39	.36	.70	.71
MZF	.68	.74	.40	.40	.57	.51	.81	.83	.61	.56	.71	.67
DZF	.51	.62	.43	.37	.71	.66	.58	.57	.42	.35	.70	.72
DOS	.53	.53	.31	.35	.67	.60	.50	.56	.38	.40	.78	.74

Note. Same rater Twins = correlation between the oldest and youngest twin, rated by M/M = mothers or F/F = fathers. Different raters Twins = cross-correlation: either oldest twin rated by mothers and youngest by fathers (M/F) or the other way around (F/M).Different raters Interparent: O = correlation between mother and father ratings for the oldest child; Y = idem for the Youngest child.

For instance, if we take for the Internalizing scale the first column "M/M": the genetic influences for boys can be estimated as $(r_{MZM} - r_{DZM}) \times 2 = (.66 - .48) \times 2 = .36$. Nonshared environmental influences for boys can be estimated as $(1 - r_{MZM}) = (1 - .66) = .34$. Following the shared environmental influences for boys can be estimated as $(2 \times r_{DZM}) - r_{MZM} = (2 \times .48) - .66 = .30$. For girls, father ratings of the Internalizing scale, and mother and father ratings of the Externalizing scale, the correlations between the MZ and DZ twin pairs can be compared in similar ways to obtain a first impression of the genetic and environmental influences.

Fitting univariate models (one for mother ratings of Internalizing, one for father ratings of Internalizing, one for mother ratings of Externalizing, and one for father ratings of

Externalizing) that estimated three factors: A, C, and E and possible sex differences, the obtained results were comparable to those calculated by comparing the MZ and DZ correlations. Take for example the Internalizing scale rated by mothers. As shown in Table 4.3, no differences between boys and girls were found. The genetic factor explained 37% of the variance, the shared environmental factor explained 32% of the variance and the nonshared environmental factor explained 31%.

Table 4.3.

Univariate estimates of genetic and environmental influences on Internalizing and Externalizing Problems rated for 10-year-old twins.

	INTERNALIZING		EXTERNALIZING	
	Mother	**Father**	**Mother**	**Father**
GENETIC	37%	♂ 69%	♂ 69%	♂ 56%
		♀ 48%	♀ 48%	♀ 51%
SHARED	32%	♂ 29%	♂ 16%	♂ 29%
		♀ 42%	♀ 33%	♀ 30%
NONSHARED	31%	♂ 32%	♂ 15%	♂ 14%
		♀ 25%	♀ 19%	♀ 18%

Different estimates for boys and girls and for mother and father ratings are found for Externalizing behavior. The sex-differences imply only a difference in the strength of the additive genetic effect and no real heterogeneity. Influences of different genes in boys and girls would be represented by lower DOS correlations in comparison to DZ correlations in same sex twins. In this study, the DOS correlation for externalizing behavior is not different from the DZ correlations.

Univariate analyses make a decomposition of the total variance in genetic, shared environmental, and nonshared environmental factors. To take rater differences into account, the information from the twin's cross-correlations has to be used. By calculating cross-correlations between mother ratings of oldest twins with father ratings of youngest twins (M/F; third column) or the other way around (F/M; fourth column), one can make a decomposition of the variance on which both kinds of raters agree. The difference between the decomposition of the variance shared between raters (i.e. common view) and the decomposition of the total variance can be used to estimate the genetic, shared environmental, and nonshared environmental influences on the variance uniquely rated by one particular rater (i.e. unique view). Take for instance for the Internalizing scale: the cross-correlations between mother ratings of oldest twins and father ratings of youngest twins (M/F) for boys. The same

comparisons between the r_{MZ} and r_{DZ} can be made to estimate the genetic influences on the variance shared by raters, namely $2 \times (r_{MZM\text{-}cross} - r_{DZM\text{-}cross}) = (.40 - .26) \times 2 = .28$. Thus we can conclude that the total genetic variance of 36% can be divided into a genetic influence for behaviors that are similarly rated by the parents of 28% and a genetic influence for behaviors that are uniquely rated by mothers of 8%. This shows that genes of the child affect the unique part of the maternal ratings, implying that the parental disagreement is not merely caused by measurement errors but that mothers in addition to the common view also assess a valid unique part of the child's behavior. Finding genetic influences for behaviors that are differently rated by mothers and fathers does not seem to be a chance finding, but arises systematically in the data. Also for the father ratings of boys and for the mother and father ratings of girls, both for the Internalizing and Externalizing scale, similar unique genetic effects are found.

To estimate the environmental influences on the variance shared by raters the interparent correlations (fifth and sixth columns for oldest and youngest twin, respectively) have to be used. Table 4.2 shows that for the Internalizing scale the interparent correlation (between mothers and fathers of the same child) in the MZM group was .64 for the oldest twin. The cross-correlation (between mothers and fathers of different children) was .40, indicating a nonshared environmental contribution on the variance shared by raters of: interparent correlation - $r_{mzm\text{-}cross} = .64 - .40 = .14$. Thus the nonshared environmental influences can be divided into an influence for behaviors that are similarly rated by both parents of 14% and an influence for behaviors that are uniquely rated by mothers of 20% (i.e. 34% - 14%). Shared environmental influences on the variance shared by raters can be estimated as $(2 \times r_{DZM}) - r_{MZM} = (2 \times .26) - .40 = .12$. Taking rater differences into account the shared environmental influences can be divided into an influence for behaviors that are similarly rated by the parents of 12% and an influence for behaviors that are differently rated by mothers of 18% (i.e. 30% - 12%). For the cross-correlations of father ratings for boys, mother and father ratings for girls, and all ratings of the Externalizing scale, similar comparisons can be made.

Rater models

As indicated by the lower χ^2 test statistic and the lower AIC in Table 4.4, the Psychometric model fitted the data better than the Rater Bias model both for the Internalizing and the Externalizing scale. This signified that although both parents partially assessed the same behaviors, there also was a component, which was unique to each rater. For sake of comparison we also performed a Cholesky or triangular decomposition (also called a Biometric model). This model can be viewed as a psychologically less informative rotation of the Psychometric model (Hewitt et al., 1992). Neither for the Internalizing scale nor for the Externalizing scale did this saturated model fit the data any better than the Psychometric model.

Table 4.4.

Model fitting statistics for Psychometric and Rater Bias model and simplification of best fitting (Psychometric) model, for 10-year-old twins' Internalizing Problems.

	INTERNALIZING						
	χ^2	df	p	AIC	χ^2	df	p
Overall model:							
Psychometric model	75.41	47	.005	-18.59			
Rater Bias model	87.54	49	.001	-10.46			
Simplification overall model:							
Factor estimates:							
No common genetic effects	132.93	49	0.00	34.93	57.52	2	.000
No unique genetic effects	97.91	51	0.00	-4.09	22.50	4	.000
No common shared env.	109.54	49	0.00	11.54	34.13	2	.000
No unique shared env.	150.94	51	0.00	48.94	75.54	4	.000
No common nonshared env.	491.78	49	0.00	393.37	416.37	2	.000
Sex differences:							
No sex dif. common effects	83.42	50	0.00	-16.58	8.02	3	.046
No sex dif. unique effects	84.37	53	0.00	-21.63	8.96	6	.176
No sex dif. common + unique	93.36	56	0.00	-18.64	**17.96**	**9**	**.036**
Rater differences:							
Unique rater effect: M- F identical	130.18	53	.000	24.18	54.78	6	.000

	EXTERNALIZING						
	χ^2	df	p	AIC	χ^2	df	p
Overall model:							
Psychometric model	55.68	47	.098	-38.32			
Rater Bias model	113.99	49	.000	15.99			
Simplification overall model:							
Factor estimates:							
No common genetic effects	361.48	49	0.00	263.5	305.80	2	.000
No unique genetic effects	128.46	51	0.00	26.46	72.78	4	.000
No common shared env.	79.86	49	0.00	-18.14	24.18	2	.000
No unique shared env.	176.25	51	0.00	74.25	120.57	4	.000
No common nonshared env.	513.55	49	0.00	415.6	457.87	2	.000
Sex differences:							
No sex dif. common effects	85.34	50	0.00	-14.66	29.66	3	.000
No sex dif. unique effects	66.35	53	0.10	-39.65	**10.68**	**6**	**.099**
No sex dif. common + unique	99.53	56	0.00	-12.47	43.86	9	.000
Rater differences:							
Unique rater effect: M- F identical	79.29	53	0.01	-26.71	23.61	6	.001

The Psychometric model was further examined for possible simplifications. None of the common and unique genetic, shared and nonshared environmental factors could be dropped from the model. Between boys and girls, the estimates of the common and the unique factors could be constrained to be equal for the Internalizing scale. For the Externalizing scale only the unique effects could be set equal for boys and girls. Mother and father ratings could not be constrained to be equal for both scales. The fit results are given in Table 4.4.

The percentages of variance explained by the common and unique genetic, shared, and nonshared environmental factors are given in Table 4.5. A major part of the variance was explained by common factors. For both the Internalizing and the Externalizing scale the largest part of the variance was explained by the common genetic factor. Common additive genetic effects explain around 30% of the variance in Internalizing behavior in boys and girls.

Table 4.5.

Genetic and environmental influences, estimated using best fitting Psychometric model, for Internalizing and Externalizing Problems rated for 10-year-old twins.

	INTERNALIZING		**EXTERNALIZING**			
	Mother	**Father**	**Mother**		**Father**	
			♂	♀	♂	♀
Genetic						
Common	26%	30%	54%	41%	56%	43%
Unique	10%	8%	11%	14%	3%	3%
Shared						
Common	18%	20%	13%	18%	13%	19%
Unique	14%	15%	7%	9%	14%	17%
Nonshared						
Common	16%	17%	8%	10%	8%	11%
Unique	16%	10%	7%	8%	6%	7%

For Externalizing behavior sex-differences in the strength of the common genetic influence is found, explaining 55% of the variance in boys and 40% of the variance in girls. The common nonshared environmental factor explained 15% of the variance for the Internalizing scale and around 10% for the Externalizing scale. The common shared environmental factor explained around 18% of the variance for both the Internalizing scale in boys and girls and Externalizing scale in girls. For Externalizing behavior in boys, only 13 % of the variance is explained by common shared environmental factors. The unique factors explained a relatively small part of the variance. For the Internalizing scale unique genetic factors explained 9%, unique shared

environmental factors explained 15%, and unique nonshared environmental explained around 15% of the variance. For the Externalizing scale unique factors also explained relatively small parts of the variance, of respectively 12% genetic influence, 8% shared, and 8% nonshared environmental influences based on mother ratings and 3% genetic influence, 15% shared, and 6% nonshared environmental influences based on father ratings.

DISCUSSION

In a sample of 2,956 Dutch 10-year-old twin pairs we studied genetic and environmental influences on Internalizing and Externalizing problems, while taking the processes underlying agreement and disagreement between maternal and paternal ratings into account.

The previous studies using CBCL data from the Netherlands Twin Register (Van der Valk *et al.*, 2001; Van der Valk *et al.*, 2002 in press) give us an unique opportunity for making comparisons and examine possible differences in genetic and environmental effects that are age related. One reason is that estimates of genetic and environmental effects are population dependent (e.g. Falconer, 1996; Plomin *et al.*, 1997). Using samples from the same country therefore excludes variation that is the result of population differences. Secondly, parameter estimates depend on the kind of raters and the methods of data collection (e.g. different questionnaires) (Van den Oord *et al.*, 2000). Because the previous studies were also based on parental ratings and used age appropriate CBCLs, comparisons are less confounded by rater and test differences. Finally, the samples used in these studies show a substantial overlap. Sampling theory (Kish, 1965) shows that differences can be more reliable observed using multiple measurements on the same sample versus using measurements in unrelated samples at different ages.

Results of this study plus similar studies in Dutch samples of 3- and 7 -year old twins, making use of the CBCL, are presented in Diagram 4.1, 4.2 and 4.3. In these diagrams the common A, C and E represent influences of additive genetic, shared environmental and unique environmental factors on aspects of the child's behavior similar assessed by both parents. The unique additive genetic effects (Unique A) represent assessments of different parts of the child's behavior for mothers and fathers. Unique shared environmental effects (Unique C) represent rater bias. Finally, unique nonshared environment (Unique E) represents measurement error.

Models for parental (dis)agreement

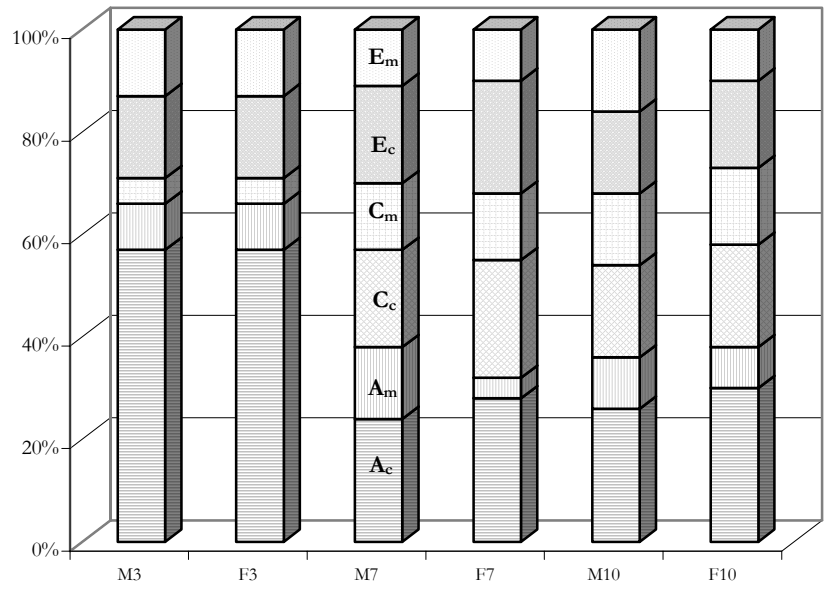
The psychometric model fitted the data better than the Rater Bias model for both scales. This implied that rater differences did not merely reflect measurement error, but were also the result of parents assessing different aspects of the child's behavior. These results are in

accordance with previous studies (Hewitt *et al.*, 1992; Van der Valk *et al.*, 2001; Van der Valk *et al.*, in press) suggesting that each parent provide additional information about the child's behavior.

The interparent correlations, representing parental agreement, can be computed by summing the A, C, and E estimates that pertain to the common part of the parental ratings.

Diagram 4.1.

Variance Decomposition of Internalizing behavior in Boys and Girls, based on Mother and Father ratings at ages 3, 7, and 10.



For Internalizing behavior a decrease in the relative importance of common effect versus unique effects is observed over the years, representing a decrease in interparent correlation, as suggested by Achenbach and colleagues (1987). At age 3 the common effects explain 73% of the total variance in Internalizing Problem behavior, while at age 10 only 64% of the total variance in Internalizing problem behavior is explained by these common factors. The better fit of the psychometric model suggests that individualization and specialization of the parent-child relation instead of a decline in the quality of parent ratings is the underlying cause of the decrease in parental agreement. For Externalizing Problem behavior less change in interparent correlation is observed. The meta-analyses by Achenbach and colleguas (1987) reported more consistency in parental agreement for undercontrolled problems (Externalizing behavior) versus overcontrolled problems (Internalizing behavior), however this was not significant for

mother/father pairs. A possible explanation, though, for the stability in parental agreement for Externalizing behavior could be that these types of behavior are better observable for an external rater than Internalizing problem behaviors and is in that manner less vulnerable to the suggested specialization or individualization of the parent-child relation.

Common aspects of Parental Ratings

The common A, C, and E factors represent the part of the child's behavior similar assessed by both parent. This part is not affected by measurement error and rater bias and represent a reliable measure of Internalizing and Externalizing problem behavior. For Internalizing behavior (Diagram 4.1) no sex-difference are found over the years. As can be seen in Diagram 4.1 the relative importance of the additive genetic effects decrease from age 3 to 7, but remain about the same from age 7 to age 10. An increase of shared environmental influences is found. At age 3 shared environmental influences are absent, while at age 7 and age 10 shared environmental influences are significant. An explanation for the presence of change between age 3 and 7 and the absence of changes between age 7 and 10 could be that the 3-7 year age interval includes children's transition to school. During this transition children must cope with many new demands like meeting academic challenges, learning school and teacher expectations, adjusting to the daily routine of a school class (Bart and Parke, 1993; Cowan *et al.*, 1994; Ladd and Price, 1987). An important aspect of this transition is the development of social relations with other children (Asher, 1990; Schneider, 1993). Although multiple pathways can be involved, poor relations with peers have shown to be a powerful predictor of behavior and emotional problems later in life.

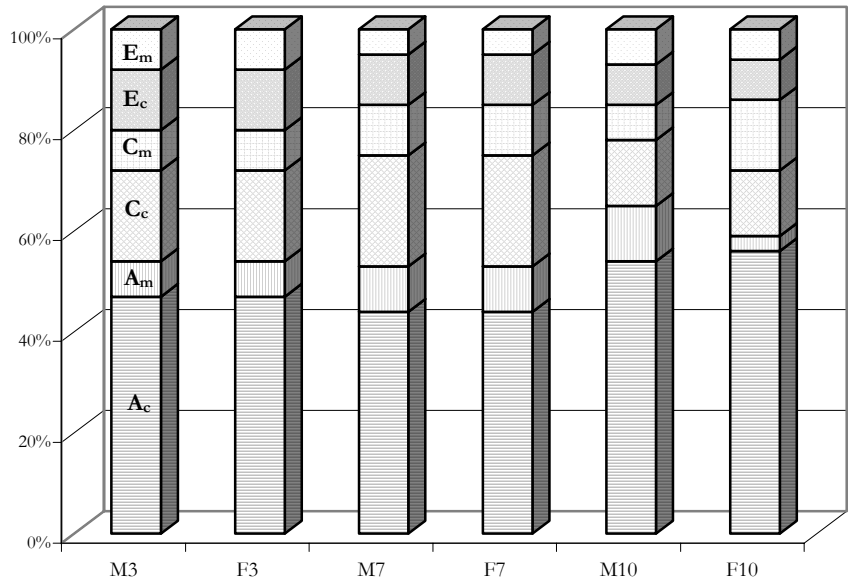
One explanation for the increase in shared environment is that if parents are only able to guide the child's behavior when he/she is able to understand other people's values and can direct its behavior accordingly, shared environmental influences are more likely to be found in older children. However, it may be important to realize that shared environment is not necessarily confined to the home environment. For instance, there are indications that these environmental effects are not merely shared by siblings but also by cousins (Van den Oord and Rowe, 1998; 1999). This suggests that shared environment reflects the wider community in which families are embedded as well (Bronfenbrenner, 1979; Parke and Kellam, 1994, p3). This point has also been stressed by Harris (1995) who argues that we should think about environmental effects in terms of group processes where peers play an important role. That is, phenomena such as within-group assimilation and between-group contrast, that increase the homogeneity of behaviors within groups and widen differences between social groups, could show as shared environment in behavior genetic analysis.

For Externalizing behavior sex difference are found at ages 7 and 10. For boys (Diagram

4.2) an increase of additive genetic effects is found going from age 7 to 10. A decrease in shared environmental influences is observed. For girls the influences of genetic and environmental factors remain stable over the years.

Diagram 4.2.

Variance Decomposition of Externalizing behavior in boys rated by mothers and fathers at ages 3, 7, and 10.



Pure nonshared environmental influences (undistorted by error or unreliability), represented by common nonshared environmental influences, were found at ages 3, 7, and 10 years. Thus idiosyncratic experiences seem to be of importance to explain both preschool and school-age children's problem behaviors.

Unique Parental Ratings

The Unique A, C, and E, explain relatively small parts of the total variance in Internalizing and Externalizing problem behavior. For Internalizing behavior, a possible specialization of the parent-child relationship over the years is represented by a relative increase of unique additive genetic factors. At age 3 the unique additive genetic factors represent 16 % of the total additive genetic effects, while at age 10 the unique additive genetic effect explain 28% of total additive genetic variance based on mother ratings and 21% of the total additive genetic variance based

on father ratings for Internalizing problem behavior. While children grow older the mother-child and father-child relation may become more distinct, because of the fact that the child's behavior becomes more diverse over the years. The diversity of behavior may create more situational specific behavior, different for mothers and fathers. The observed drop of the unique additive genetic effects in father ratings, representing the unique view of the father, from age 3 to 7 is recovered going from age 7 to age 10. This drop in additive genetic influences is not observed in the common factor, which could suggest that the drop in unique additive effects for the father ratings is the result of sample fluctuations, rather than a real change in behavior. Future studies are necessary to accept or reject this finding.

For Externalizing behavior less change is observed. The relative importance of the unique additive genetic effects based on mother ratings, representing the unique view of the mother on the boys' behavior, seems to be relatively stable over the years. Father's unique view, however, represented by the unique additive genetic effect in father ratings, shows a remarkable drop at age 10. The same pattern of unique additive genetic influences are found in girls' behavioral ratings, presented in Diagram 4.3. Fathers seem to add no unique view on Externalizing behavior at age 10 in both boys and girls. This could again be the result of sample fluctuations and future studies in this sample may be helpful to determine whether this effect will persist at older ages.

Rater bias was included in the estimate of the unique shared environmental factor, accounting for at most 17% of the variance for both the Internalizing and Externalizing scale. For Internalizing behavior an increase in Rater bias could be observed from age 3 to 7, but again this is stabilized from age 7 to age 10. The fact that this increase occurs in both the unique as well as the common aspect of shared environment suggests a real change in behavior due to the development of the child. If it the increase represented rater bias solely, only an increase in unique shared environmental influences would be observed. For Externalizing behavior the influence of rater bias remained constant.

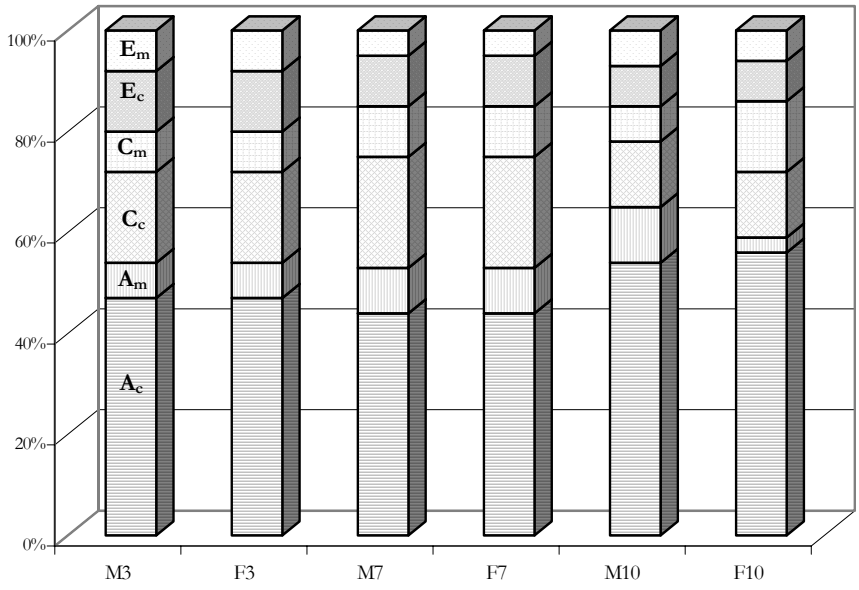
Measurement errors and unreliability were estimated in the unique nonshared environmental factor. However, neither for the Externalizing scale nor for the Internalizing scale did this factor account for more than 11% of the variance, except Internalizing behavior rated by the mother (16%).

At age 10, a trend that started at age 7 seems to be continued. At age 3, for mother and fathers ratings, the estimates of unique factors were allowed to be constrained equal for both Internalizing and Externalizing behavior. At age 7 however, the estimates of unique factors for Internalizing behavior should be considered different for mothers and fathers. At age 10 estimates of the unique factors for both Internalizing and Externalizing behavior could not be constrained equal for mother and fathers. This result indicates that different amounts of the

total variance are influenced by the 'unique view' of a parent. This again might be a representation of specialization in the relation parents do have with their children. Parents may be observing more situational specific behavior with an increase of 'movement space' of their children when they grow older, resulting in more distinct unique views for mothers versus fathers.

Diagram 4.3.

Variance Decomposition for Externalizing behavior in Girls rated by mothers and fathers at ages 3, 7, and 10.



Limitations

Results of this study indicate that parental ratings are a valuable instrument for assessing behavioral and emotional problems in children. Using both mother and father ratings will give more reliable results by decreasing measurement error and rater bias for the part of the behavior similar assessed by mothers and fathers. Further, results of these studies indicate that parents assess an unique aspect of their child's behavior so that the combination of mother and father ratings will give a more complete picture of the child's behavior. Although parents have the advantage that they observe their children over longer periods of time and can witness both frequent and rare behaviors, they mainly interact with their children in the home environment. Adding raters such as teacher who observe the child's behavior in other situations may contribute valuable information. For instance, comparison of the predictive power of parent and teacher information showed that teacher scores were a stronger

prediction of poor outcomes than parents (Verhulst *et al.*, 1994). Although teachers report fewer problems than parents about the same children, their reports apparently are informative with respect to later functioning. Further, self-reports might be valuable as well. Children may behave in a different manner when they are with their parents or their teacher. Parents and teachers can only rate those aspects of their children's behavior of which they are aware. Children, though, may be engaged in a variety of behaviors about which they do not tell their parents or teacher. Obviously, self-report becomes more important with increasing age.

The best selection of raters may depend on the type of problems that are studied. There is considerable evidence that parents are more likely to report symptoms of overactivity, inattention, and oppositional behavior than their children (Edelbrock *et al.*, 1986; Herjanie and Reich, 1982; Kashani *et al.*, 1985; Loeber *et al.*, 1991). On the other hand, children more frequently endorse emotional symptoms, including phobias and obsessional behavior (Herjanic and Reich, 1982) and depression (Angold *et al.*, 1987; Kashani *et al.*, 1985). Further, Loeber and colleagues (1989, 1991) have argued that parents and teachers are better informants for hyperactivity and oppositional behavior, while children and parents should be used to elicit conduct disorder symptomatology.

Psychopathology in parents seems to be correlated. Significant spousal correlations are found for more Internalizing behaviors like depression and anxiety as well as Externalizing behaviors such as antisocial behavior (Stallings et al., 1997; Krueger *et al.*, 1998; Dufouil and Alperovitch, 2000; Mathews and Reus, 2001). These correlations could be a result either from assortative mating or contagion/interaction effects. For assortative mating, nonrandom mating occurs based on the psychopathology in both parents and in that case is a matter of selection. Contagion effects arise after mating and could be a result of the length of the relationship.

Assortative mating is important for genetic research for two reasons. First, assortative mating increases genetic variance in a population (Falconer and Mackay, 1996). In other words, positive assortative mating increases variance in that the offspring differ more from the average than they would if mating were random. Even though spouse correlations are modest, assortative mating can greatly increase genetic variability in a population, because its effects accumulate generation after generation. Assortative mating is also important because it affects estimates of heritability. Positive assortative mating increases the resemblance between fraternal or dizygotic twins because it renders the parents of these twins more similar than they would be if there were no assortments. Identical or monozygotic twins, however, are already at the point of maximum genetic resemblance, and are thus unaffected by positive assortative mating (Fulker, 1988). This will result in an overestimation of shared environmental influences and an underestimation of additive genetic effects. A parent-offspring design would be necessary to investigate whether influences of shared environment are overestimated due to

assortative mating.

Another effect of parental resemblance in psychopathology is that shared environmental effects of the part of the child's behavior assessed by both parents can be overestimated. Several studies suggest that depression in mothers may lead to their overestimating their children's symptomology (Fergusson and Horwood, 1987). In one study (Breslau *et al.*, 1988), mothers who were depressed rated their children as showing a greater number of symptoms of all psychiatric syndromes. Like mothers, fathers' reports of their children's behavioral problems are influenced by their own level of psychological symptoms (Phares *et al.*, 1989; Jensen *et al.*, 1988). The consequence of the facts that a) parents tend to have similar levels of psychopathology and b) levels of parental psychopathology affect ratings of problem behavior in their children is that the rater bias components of mothers and fathers become correlated. Because this shared rater bias component will effect MZ and DZ twin correlations in the same way, it will show as shared environmental effects on the common part of the parental ratings. The inclusion of measures of parental psychopathology or the use of different type of raters such as teachers will be helpful to account for these correlated rater bias effects.

In summary, besides parental agreement, unique parental views on their children's behaviors seem to be significant at age 10. These results are in line with the findings in comparable samples of Dutch twins at ages 3 and 7 years. Additive genetic factors remain important as a source of individual differences in Internalizing and Externalizing problem behavior. Shared environmental influences, however, are also substantial. The changes in genetic and environmental effects, occurs mainly between age 3 and 7. No major changes are observed between age 7 and age 10. The significant influences of additive genetic factors indicate an innate vulnerability to childhood psychopathology. The influences of nonshared environmental influences suggest the importance of pure idiosyncratic experiences.

5

When mom's and dad's do and don't agree: A study of parent ratings of Internalizing and Externalizing problem behavior in 12-year-old twins

---

This chapter is submitted as:

Bartels, M., Hudziak, J.J., Boomsma, D.I., Rietveld, M.J.H., Van Beijsterveld, C.E.M., and Van den Oord, E.J.C.G. (2002). When mom's and dad's do and don't agree: A study of parent ratings of Internalizing and Externalizing problem behavior in 12-year-old twins. *Journal of the American Academy of Child and Adolescent Psychiatry*, submitted.

ABSTRACT

Previous studies on parent reports of 3, 7, and 10-year-old twins Internalizing and Externalizing problems have emphasized the importance of understanding the source and impact of agreement and disagreement between maternal and paternal ratings, when estimating the genetic and environmental contributions to these behaviors. In reports on younger twins, a Psychometric model, which assumes that each parent assesses an unique aspect of their child's behavior, provided the best explanation for differences in parental ratings. The purpose of this study was to investigate if a similar explanation for disagreement between parents can be supported in older twins in the peri-pubertal (12-year-old) age group. Further, genetic and environmental influences on Internalizing and Externalizing problem behavior are estimated. Child Behavior Checklists filled in by mothers and fathers, were collected for a sample of 1481 Dutch 12-year-old twin pairs. Genetic and environmental influences on parental reports of Internalizing and Externalizing syndromes was examined using models that corrected for possible rater bias, unique rater effects and unreliability. A Psychometric model (one that posits that parents partly assess unique aspects of their child's behavior) fitted the data of both scales significantly better than a Rater Bias model (one that posits that disagreement is bias or unreliability). Common factors (influencing behaviors similarly assessed by both parents) were more important than unique factors (influencing behaviors uniquely assessed by one parent). Significant influences of additive genetic, shared environmental and unique environmental factors have been found for Internalizing and Externalizing syndromes. Rater bias and unreliability were included in the estimates of the unique factors, which were small. The best fitting model is one that implies that disagreement between the parents is due to the fact that mom and dad are providing information from their own perspective. The significant influences of additive genetic factors indicate a possible innate vulnerability to childhood psychopathology. The influences of common nonshared environmental influences suggest the importance of pure idiosyncratic experiences. Significant influences of environment shared by both members of a twin pair are represented by the common shared environmental factor. Finally, the Psychometric model argues for inclusion of information from both parents when assessing these common behaviors as it is clear that mothers and fathers provide a unique perspective on their children.

INTRODUCTION

Behavioral/emotional problems are common among children. Two prevalence studies in representative samples of Dutch children reported that 7-8% of preschool and school-aged children show problem behavior (Verhulst *et al.*, 1985a; 1985b; 1989; 1997; Verhulst and Akkerhuis, 1986). These studies, further, indicate that Internalizing problems (anxious/depressed behavior, withdrawn behavior) are more prevalent in girls, especially in adolescence, and Externalizing problems (aggressive behavior, rule breaking behavior) are more prevalent in boys. Most problem behaviors in children are considered to be influenced by multiple genes and environmental influences. In this respect different kinds of behaviors, such as Internalizing or Externalizing behavior, generally do not fall in distinct categories of behavior that are either present or absent, but involve quantitative variations of behaviors that most children display to some degree.

One approach to quantify children's problem behavior is by asking the parents to score behavioral and emotional problems on behavioral questionnaires. Traditionally, data from mothers have been used to determine whether or not a child has a psychiatric syndrome. Information from fathers has been difficult to obtain, and once obtained, difficult to interpret, given the often low levels of agreement between mother and father reports. A meta-analysis by Achenbach, McConaughy, and Howell (1987) showed a mean correlation of .60 between maternal and paternal ratings of the same child. This underscores that parents are able to assess their child's behavior, for if parental ratings would reflect nothing but error the correlation between their ratings would probably be low. However, a high correlation does not necessarily imply that parents are assessing the same phenotype in their children (Hewitt *et al.*, 1992). The correlations may be high even when the parents are assessing different behaviors in their children, because the parental correlation may predominate over the variance specific to a given parent. Conversely, different forms of rater bias and unreliability may lower the correlation between parents even though parents may be assessing exactly the same phenotype in their children. From a clinical point of view, it remains a struggle to determine what to do with the disagreement. Is it best to assume that there is 'one best informant', that one parent is 'more reliable than the other', or that parents present a unique viewpoint on his or her child, thus providing unique and valuable information to be used in assessment? Accordingly, it is difficult to draw firm conclusions about the processes underlying the (dis)agreement between parental ratings on the basis of correlations alone.

To study agreement and disagreement between parental ratings, Hewitt and colleagues (1992) proposed so-called Rater Bias and Psychometric models that combine data of mothers and fathers and can be estimated using genetically informative data. The Rater Bias model assumes that parents assess exactly the same behaviors in the child (common behavioral view) and that they share a common understanding of the behavioral descriptions. This may apply

when both parents are equally confronted with the behaviors shown by the child (for instance at home). Disagreement between the raters is regarded as error, resulting from rater bias and/or unreliability. Sources of rater bias are stereotyping, employing different normative standards, or having certain response styles, i.e. judging problem behaviors more or less severely. Because these types of bias may differ between raters, they may also lead to disagreement between raters. Unreliability can become an important source of disagreement when raters cannot give an accurate description about relevant behaviors. For instance, evidence is found that parents may be relatively insensitive to affective disturbances in children (Angold *et al.*, 1987).

The Psychometric model assumes that each parent assesses an unique aspect of his or her child's behavior. This will occur when the parent also observes the child in distinct situations where they are exposed to distinct samples of the behavior (unique behavioral view). For instance, the parent who usually brings the child to school may also be more familiar with the child's behavior outside the home. Moreover, each parent may interact differently with the child (Achenbach *et al.*, 1987). These unique interactions between a parent and a child may allow each parent to provide additional information about the child's behavior, apart from the information on which they both agree. Disagreement in this model does not merely arise due to unreliability and/or rater bias, but also because each parent contributes, from his own perspective, different but valid information on the child's functioning. The psychometric model tests this by examining whether there are significant child genetic effects on the parents' unique behavioral views. If the behaviors uniquely rated by the parents are shown to be influenced by genetic factors of the child, the parent must have been assessing 'real' unique behavioral views. For error and/or unreliability cannot cause the systematic effects necessary for the model to estimate genetic influences.

Several quantitative genetic studies have used the Child Behavior Checklist (CBCL) (Achenbach, 1991; 1992) to examine the etiology of children's problem behaviors (Edelbrock *et al.*, 1995; Gjone and Stevenson, 1997; Hudziak *et al.*, 2000; Leve *et al.*, 1998; Schmitz *et al.*, 1995; Silberg *et al.*, 1994; Van den Oord *et al.*, 1996; Van der Valk *et al.*, 1998a; Van der Valk *et al.*, 1998b; Zahn-Waxler *et al.*, 1996). Yet, only a few studies employed models that incorporated rater differences. Rowe and Kandel (1997) collected the CBCL completed by mothers and fathers for their oldest two offspring (aged 9 to 17) in 76 families. They did not fit either Psychometric or Rater Bias models. Still, their results, based on an 'individual view-shared view' model, showed that the parental ratings contained a substantial shared behavioral view. Simonoff *et al.* (1995), in a study of 282 twin pairs aged 8 to 16, also found evidence in favor of a shared behavioral view for antisocial behaviors. However, from their analyses they could not determine what underlay the shared parental view and described it as due to a shared set of expectations of the parents against which both twins were rated. Hewitt *et al.* (1992)

applied both the Rater Bias and Psychometric model on parental ratings of the Internalizing scale (CBCL) for 983 twin pairs. They found that both for their prepubertal cohort (8 to 11 years) and for their pubertal cohort (12 to 16 years) the Psychometric model fitted the data better than the Rater Bias model. Hewitt and colleagues concluded that for the Internalizing scale, mothers and fathers rate the same phenotype in their children (i.e. have a shared behavioral view). Unique genetic influences were found, implying that the rater differences reflected the existence of real unique behavioral views and not just error and bias. Further insight in issues of rater bias is presented by van der Valk and colleagues (2001; 2002). Rater bias models and Psychometric models were tested on a large groups of 3-and 7-year-old Dutch twins. As in the previous studies, the Psychometric model fitted the data significantly better at both ages. In a twin sample of 10-year-old twins the same results are found (Bartels *et al.*, submitted). This again indicates 'real' unique behavioral views for mother and fathers separately, beside a common view on behavior for mothers and fathers.

In the present study Rater Bias and Psychometric models were fitted to the Internalizing and Externalizing scale of 1481 Dutch 12-year-old twin pairs to examine whether the results from previous studies could be confirmed in a sample of 12-year-old twins. The analyses will determine if there are differences in how parents report on Internalizing and Externalizing behavior in 12-year-olds versus 3, 7, and 10-year-olds. Tests for differences in how mothers and fathers report on Internalizing behavior and on Externalizing behavior will be conducted. We will test for informant by gender differences to determine if mothers report differently on their daughters or sons. In short, the processes underlying parental disagreement were examined in a sample of 12-year-old twin pairs and the etiology of Internalizing and Externalizing Problems was studied.

METHODS

Subjects

All participants are registered by the Netherlands Twin Registry (NTR), kept by the Department of Biological Psychology at the Vrije Universiteit in Amsterdam. Of all multiple births in the Netherlands, 40-50% is registered by the NTR (Boomsma *et al.*, 1992; Boomsma, 1998; Boomsma *et al.*, 2002). For this study, data from twins from the birth cohorts 1986 - 1990 were used. Questionnaires were mailed to families within three months of the twins' twelfth birthday. After two to three months reminders were sent and four months after the initial mailing persistent non-responders were contacted by phone. Families whose address was no longer available were included in the nonresponse group. From the original sample 80 twin pairs were excluded because either one or both of the children had a disease or handicap that interfered severely with daily functioning at age 12 or at a younger age. The final sample for analysis consists of 1481 mother ratings and 1156 father ratings.

Zygosity was determined for 472 same-sex twin pairs by DNA analyses or blood group polymorphisms. For all other same-sex twin pairs zygosity was determined by discriminant analysis, using questionnaire items. Parents were asked how much the twins resembled each other in facial structure, hair color, facial color, eye color, and whether they were ever mistaken for each other by the parents themselves, by family, or by strangers. They were also asked if the twins were as much alike as two peas in a pod, whether it was difficult for the parents to separate the twins on a recent picture, and how similar the twins' hair structure was (For details see Rietveld *et al.*, 2000).

This left a sample of 283 monozygotic males (MZM), 231 dizygotic males (DZM), 315 monozygotic females (MZF), 228 dizygotic females (DZF), and 424 dizygotic opposite sex (DOS) twin pairs. In general, mothers' response rate outnumbers fathers' response rate. Therefore the data could be further divided into twin pairs for which both mother and father had replied (225 MZM, 180 DZM, 240 MZF, 187 DZF, and 324 DOS) and twin pairs for which only mothers had replied (58 MZM, 51 DZM, 75 MZF, 41 DZF, and 100 DOS).

Measures

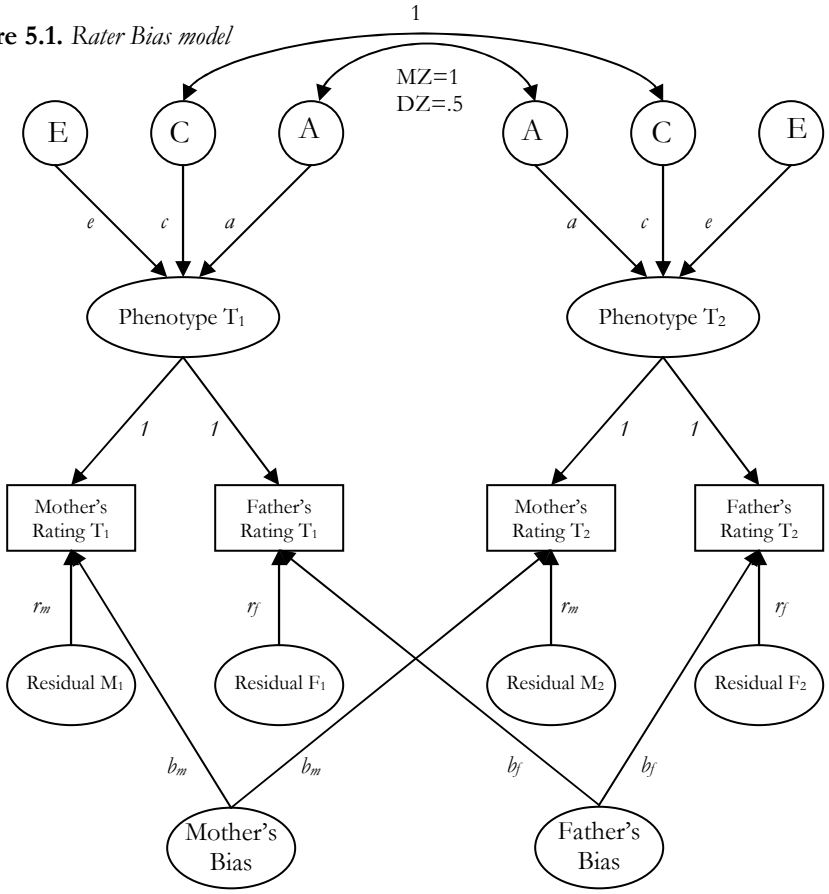
The Child Behavior Checklist (CBCL 4-18) (Achenbach, 1992) was developed for parents to score the behavioral and emotional problems of their 4-to-18-year-old children. It consists of 120 problem items that are scored by the parents on a 3-point scale based on the occurrence of the behavior during the preceding 6 months: 0 if the problem item was not true of the child, 1 if the item was somewhat or sometimes true, and 2 if it was very true or often true. The syndrome scales were composed according to the 1991 profile (Achenbach, 1991). Dutch syndrome scales and comparability with the syndrome scales as developed by Achenbach are reported in the Dutch manual (Verhulst *et al.*, 1996). In this manual the two broad-band scales Internalizing and Externalizing are analyzed. The Internalizing scale consists of the Anxious/Depressed, Somatic Complaints and Withdrawn subscales. The Externalizing scale consists of the Aggressive and Rule Breaking Behavior subscales. For the Internalizing scale subjects were only included if not more than 3 items were missing for the Anxious/Depressed scale, not more than 2 items were missing for Somatic Complaints and Withdrawn scales. For the Externalizing scale the inclusion criterion was not more than 3 items missing for the Aggressive and Rule Breaking Behavior scales. This ensured that the two syndrome scales were always composed of all problem behaviors loading on that scale.

The data were square root transformed to approximate normal distributions that are required for maximum likelihood estimation. After transformation, all skewness and kurtosis indices were between -1.0 and 1.0, implying that not much distortion is to be expected (Muthén and Kaplan, 1985).

Structural equation modeling of twin data rated by more than one rater

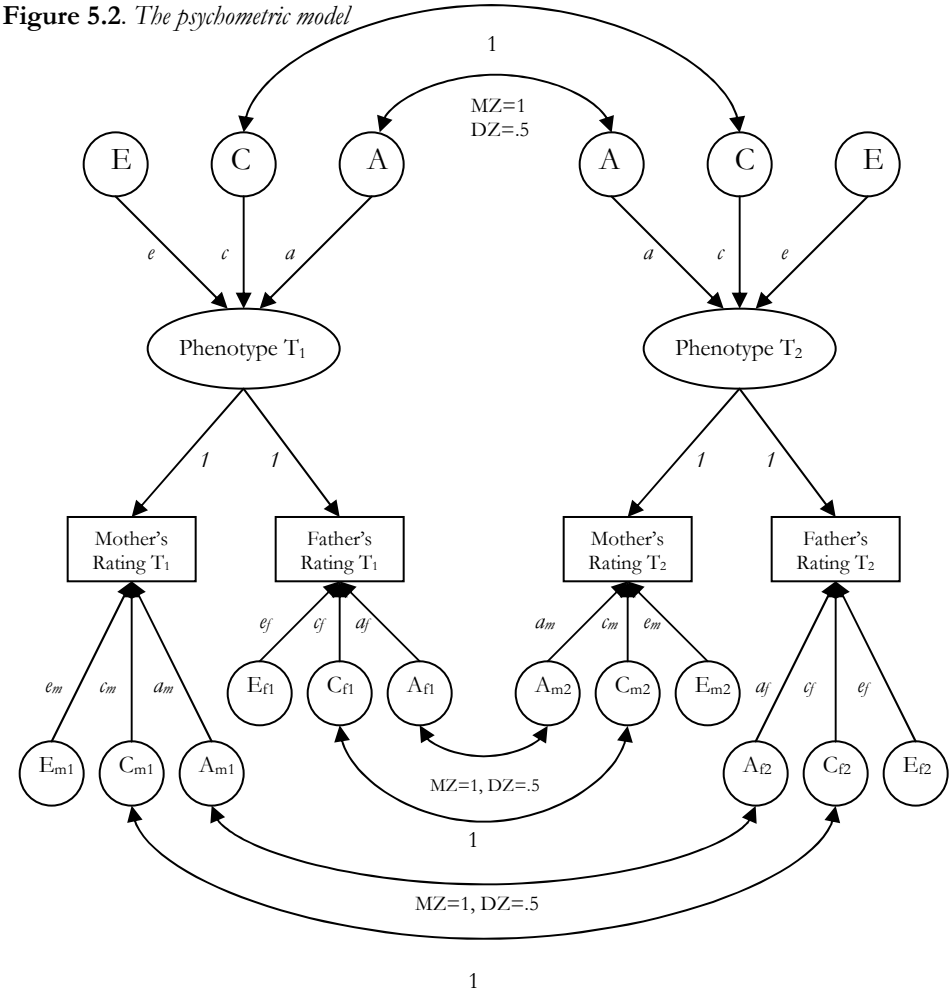
Standard approaches to structural equation modeling of twin data are discussed in detail elsewhere (Bartels *et al.*, submitted ; Van der Valk *et al.*, 2001; 2002). In this paper we present a discussion of the Rater bias and the Psychometric models.

Figure 5.1. *Rater Bias model*



In the Rater Bias model (Hewitt *et al.*, 1992) (Figure 5.1) the phenotypes of the twins are a function of three common factors underlying the ratings of both mothers and fathers: a genetic factor (A), a shared environmental factor (C), and a nonshared environmental factor (E). In addition to these three common factors unique factors are modeled: a maternal rater bias factor, a paternal rater bias factor, and residual (unreliability) factors affecting each rater. The influence of the common factors is assumed to be independent of the maternal and paternal rater bias and unreliability factors.

Figure 5.2. *The psychometric model*



The Psychometric model (Hewitt *et al.*, 1992) (Figure 5.2) also estimates the influence of a genetic (A), a shared environmental (C), and a nonshared environmental factor (E) common to the phenotypes of the twins as rated by both parents. In addition three unique factors, a genetic ($A_{m/f}$), shared environmental ($C_{m/f}$), and nonshared environmental factor ($E_{m/f}$) are estimated for the ratings of mother/father. Disagreement between parents in this model can be caused by unique behavioral views, leading to different but valid information of each rater. These unique behavioral views can have their own unique influences, estimated in the unique additive genetic factors. Disagreements can also be caused by rater bias, which will confound the unique shared environmental effects, or by unreliability which will confound the unique nonshared environmental effects. The three common factors loading on the twins' phenotypes contain only reliable variance, causing the common nonshared environmental factor to contain

only pure independent environmental effects (McArdle and Goldsmith, 1990) and the common shared environmental factor to contain only pure shared environmental effects.

Model fitting

The program Mx (Neale *et al.*, 1999) was used to analyze the data through a simultaneous analysis of the 4 × 4 variance-covariance matrices in the five zygosity by sex twin groups (MZM, DZM, MZF, DZF, DOS) where both mother and father ratings were available, and the 2 × 2 variance-covariance matrices in the five zygosity by sex twin groups with only mother ratings. The model describes the observed variance-covariance matrices adequately when the residual variance-covariance matrices are trivially small. A good model is indicated by a low non-significant χ^2 test statistic ($P > .05$). Apart from the χ^2 test statistic, Akaike's Information Criterion (AIC = χ^2 - 2 × degrees of freedom) was computed. The lower the AIC the better the fit of the model to the observed data.

Fitting the Rater Bias and Psychometric model to the data showed which model described the processes involved in either agreement or disagreement between the parental ratings best. Monozygotic twin covariances and dizygotic twin covariances were modeled, assuming a correlation between the twins' shared environmental factors of 1.0, regardless of twin type, and a genotypic correlation of 1.0 for monozygotic twins and 0.5 for dizygotic twins. For dizygotic twins of opposite sex a covariance lower than .5 can occur when sex-specific genes influences behavior in boys or girls, so-called heterogeneity.

Table 5.1.

Means (standard deviations) and sample sizes for the Internalizing and Externalizing scale, in a 12-year-old twin (per zygosity) and a 4-11 year-old Dutch community sample.

		INTERNALIZING			EXTERNALIZING		
		Mother	**Father**	**N (M/F)**	**Mother**	**Father**	**N (M/F)**
♂	**MZM**	4.11(4.63)	3.20 (4.40)	561/447	7.25 (6.91)	6.07 (6.84)	569/447
	DZM	4.09 (4.67)	3.49 (4.50)	453/362	6.71 (6.52)	5.94 (6.52)	468/363
	DOS	3.79 (4.68)	2.75 (3.49)	410/317	6.70 (6.76)	5.53 (5.62)	416/318
	COM	5.36 (5.36)		440	6.35 (6.13)		440
♀	**MZF**	4.79 (4.82)	3.08 (3.26)	617/469	5.18 (5.28)	4.10 (4.58)	629/479
	DZF	4.63 (4.55)	4.05 (4.65)	448/377	4.89 (5.04)	4.19 (4.62)	459/377
	DOS	4.31 (4.43)	2.98 (3.35	427/334	4.79 (5.30)	3.62 (4.09)	440/333
	COM	6.32 (5.93)		456	5.21 (5.43)		456

Note. MZM/DZM = Monozygotic/Dizygotic males, MZF/DZF = Monozygotic/Dizygotic females, DOS = Dizygotic opposite sex, COM= Dutch community sample, N (M/F) = number of children for Mothers (M) and Fathers (F).

A first indication for the significance of heterogeneity is a lower twin correlation for dizygotic twins of opposite sex in comparison to same sex dizygotic twins. Based on the twin correlations no test for heterogeneity is conducted. The strength of the common and unique genetic and environmental influences, though, were allowed to differ for boys and girls. This model was further examined for possible simplifications. It was tested whether the common and/or unique factors could removed from the model, whether estimates for boys and girls could be constrained to be the same, and if the unique factors for mothers and fathers could be constrained to be equal. The only factor that was never dropped from the model was the unique nonshared environmental factor, because apart from the influences of idiosyncratic experiences, measurement errors are also estimated in this factor.

RESULTS

Description of the data

The untransformed mean problem scores and standard deviations of the twin sample and those of a Dutch community sample of 4-11-year-old children (Verhulst *et al.*, 1996) are given in Table 5.1.

TABLE 5.2.

Correlations (ratings given by the same rater), and cross-correlations (ratings given by different raters) between the twins and the interparent correlations, per zygosity, for 12-year-olds.

	INTERNALIZING						EXTERNALIZING					
	same rater		different rater				same rater		different rater			
	TWINS		TWINS		INTER PARENT		TWINS		TWINS		INTER PARENT	
	M/M	F/F	M/F	F/M	O	Y	M/M	F/F	M/F	F/M	O	Y
MZM	.73	.74	.48	.47	.66	.60	.87	.90	.71	.68	.78	.75
DZM	.46	.47	.31	.36	.67	.68	.54	.61	.39	.39	.71	.71
MZF	.73	.69	.41	.43	.55	.61	.84	.83	.66	.62	.70	.72
DZF	.54	.61	.38	.40	.65	.64	.56	.58	.44	.35	.74	.69
DOS	.59	.49	.41	.32	.63	.61	.61	.57	.42	.48	.74	.69

Note. Same rater Twins = correlation between the oldest and youngest twin, rated by M/M = mothers or F/F = fathers. Different raters Twins = cross-correlation: either oldest twin rated by mothers and youngest by fathers (M/F) or the other way around (F/M).Different raters Interparent: O = correlation between mother and father ratings for the oldest child; Y = idem for the Youngest child.

For both the Externalizing and the Internalizing scale, the ratings given to the twins were quite similar to the ratings given to the Dutch community sample. Within the twin group, one-way ANOVA showed no significant mean differences between MZ and DZ twin pairs for boys (MZM vs. DZM) or for girls (MZF vs. DZF), neither for maternal nor for paternal ratings. Comparing boys and girls (MZM vs. MZF, and DZM vs. DZF), both mothers and fathers gave significantly higher ratings to the boys for the Externalizing scale (MZ: Mothers: $F_{(1,1196)}$=34.421, p=.000; Fathers: $F_{(1,924)}$=26.960, p=.000; DZ: Mothers: $F_{(1,925)}$=22.505, p=.000; Fathers: $F_{(1,738)}$=17.853, p=.000). For the Internalizing scale ratings MZ female twins were rated higher by mothers in comparison to DZ female twins ($F_{(1,1176)}$=6.072, p=.014). Comparing mother and father ratings, a paired T-Test showed that the ratings for the both the Internalizing and Externalizing scales given by mothers were significantly higher than ratings given by fathers for both boys and girls (Internalizing: boys: T=7.566, df=1093, p=.000; girls: T=9.744, df=1139, p=.000; Externalizing: boys: T=6.729, df=1113, p=.000; girls: T=7.339, df=1174, p=.000). Thus, MZ and DZ twin pairs were not rated differently, allowing to use the twin data for genetic analyses. Boys did receive higher ratings than girls for the Externalizing scale. For this same scale and for the Internalizing scale, mothers gave higher ratings to their twin children than fathers did, implying possible rater differences.

The homogeneity of the variance was tested in Mx. No differences could be found in the variances of MZM, DZM, MZF, DZF, and DOS, for both scales.

Univariate analysis

Table 5.2 shows, for both the Internalizing and Externalizing scale, in the first and second column the correlations between the twins rated by the same rater (mother *or* father rated both children), and in the third and fourth column the cross-correlations between the twins each rated by a different rater (mother and father each rated one child). In the fifth and sixth column the interparent correlations between mothers and fathers are given, both for oldest and youngest twin. The interparent correlations were comparable for both oldest and youngest twin for all zygosity by sex groups. On average, the interparent correlations for the Internalizing scale were .63, and for the Externalizing scale .72. This resembled the interparent correlations obtained in the Dutch norm group (Verhulst *et al.*, 1996).

Fitting univariate models (one for mother ratings of Internalizing, one for father ratings of Internalizing, one for mother ratings of Externalizing, and one for father ratings of Externalizing) estimated three factors: A, C, and E and possible sex differences (Table 5.3). Significant influences of additive genetic, shared environmental and unique environmental influences are found for both Internalizing and Externalizing behavior. Different estimates for boys and girls and for mother and father ratings are found for Internalizing and Externalizing behavior. The sex-differences imply only a difference in the strength of the additive genetic

effect and no real heterogeneity. Influences of different genes in boys and girls would be represented by lower DOS correlations in comparison to DZ correlations in same sex twins. In this study, the DOS correlations for Externalizing behavior are not different form the DZ correlations.

Table 5.3.

Univariate estimates of genetic and environmental influences on Internalizing and Externalizing Problems rated for 12-year-old twins.

	INTERNALIZING		EXTERNALIZING	
	Mother	**Father**	**Mother**	**Father**
GENETIC	36%	♂ 50%	♂ 57%	♂ 55%
		♀ 34%	♀ 52%	♀ 51%
SHARED	37%	♂ 23%	♂ 30%	♂ 34%
		♀ 38%	♀ 32%	♀ 32%
NONSHARED	27%	♂ 27%	♂ 13%	♂ 11%
		♀ 28%	♀ 16%	♀ 17%

Rater models

As indicated by the lower χ^2 test statistic and the lower AIC in Table 5.4, the Psychometric model fitted the data better than the Rater Bias model both for the Internalizing and the Externalizing scale. This finding signifies that although both parents partially assessed the same behaviors, there also was a component, which was unique to each rater. For sake of comparison we also performed a Cholesky or triangular decomposition (also called a Biometric model). This model can be viewed as a psychologically less informative rotation of the Psychometric model (Hewitt *et al.*, 1992). Neither for the Internalizing scale nor for the Externalizing scale did this saturated model fit the data any better than the Psychometric model.

The Psychometric model was further examined for possible simplifications. Non of the common and unique genetic, shared and nonshared environmental factors could be dropped from the model. Between boys and girls, the estimates of the common and the unique factors could be constrained to be equal for the Internalizing scale. For the Externalizing scale only the unique effects could be set equal for boys and girls. Mother and father ratings could be constrained to be equal for both scales. The fit of the best model is given in Table 5.4. The percentages of variance explained by the common and unique genetic, shared, and nonshared environmental factors are given in Table 5.5. A major part of the variance was explained by common factors. For the Externalizing scale the largest part of the variance was explained by

the common genetic factor and sex-differences in the strength of this common genetic influence is found, explaining 48% of the variance in boys and 40% of the variance in girls. For the Internalizing scale, 22% of the variance is explained by common genetic factors. The common nonshared environmental factor explained 16% of the variance for the Internalizing scale and around 8% for the Externalizing scale. The common shared environmental factor accounted for around 25% of the variance for the Internalizing scale in boys and girls and 20% of the variance in the Externalizing variance in both boys and girls. The unique factors explained a relatively smaller part of the variance. For the Internalizing scale unique genetic factors explained 15%, unique shared environmental factors explained 11%, and unique nonshared environmental explained around 11% of the variance. For the Externalizing scale unique factors also explained relatively smaller parts of the variance respectively 9% additive genetic, 10% shared environmental, and 6% nonshared environmental influences for boys and 11% additive genetic, 13% shared environmental, and 8% nonshared environmental influences for girls.

Table 5.4.

Model fitting statistics for Psychometric and Rater Bias model and simplification of best fitting (Psychometric) model, for 12-year-old twins' Internalizing Problems.

	INTERNALIZING						
	χ^2	**df**	**p**	**AIC**	**χ^2**	**df**	**p**
Overall model:							
Psychometric model	51.557	47	0.30	-42.44			
Rater Bias model	78.467	49	0.01	-19.53			
Simplification overall model:							
Factor estimates:							
No common genetic effects	73.595	49	.01	-24.60	22.038	2	.000
No unique genetic effects	86.08	51	0.00	-15.92	34.523	4	.000
No common shared env.	80.224	49	0.00	-17.78	28.667	2	.000
No unique shared env.	75.905	51	0.01	-26.10	24.348	4	.000
No common nonshared env.	257.69	49	0.00	159.69	206.133	2	.000
Sex differences:							
No sex dif. common effects	56.771	50	0.24	-43.23	**5.214**	3	**.157**
No sex dif. unique effects	54.122	53	0.43	-51.88	**2.565**	6	**.861**
No sex dif. common + unique	60.47	56	0.32	-51.53	**8.931**	**9**	**.445**
Rater differences:							
Unique rater effect: M- F identical	55.654	53	0.38	-50.35	4.097	6	.000
Simplified model	**63.981**	**59**	**0.31**	**-54.02**			

Table 5.4.- continued

Model fitting statistics for Psychometric and Rater Bias model and simplification of best fitting (Psychometric) model, for 12-year-old twins' Externalizing Problems.

	EXTERNALIZING						
	χ^2	df	p	AIC	χ^2	df	p
Overall model:							
Psychometric model	38.532	47	0.81	-55.47			
Rater Bias model	74.875	49	0.01	.23-13			
Simplification overall model:							
Factor estimates:							
No common genetic effects	199.621	49	0.00	101.62	161.089	2	.000
No unique genetic effects	80.345	51	0.00	-21.66	41.813	4	.000
No common shared env.	59.763	49	0.14	-38.24	21.231	2	.000
No unique shared env.	91.164	51	0.00	-1.084	52.632	4	.000
No common nonshared env.	186.068	49	0.00	88.07	147.536	2	.000
Sex differences:							
No sex dif. common effects	57.103	50	0.23	-42.90	18.571	3	.000
No sex dif. unique effects	48.232	53	0.66	-57.77	**9.700**	6	**.138**
No sex dif. common + unique	71.068	56	0.09	-40.93	32.536	**9**	.000
Rater differences:							
Unique rater effect: M- F identical	50.524	53	0.57	-55.48	**11.992**	6	.062
Simplified model	**55.070**	**56**	**.051**	**-56.93**			

DISCUSSION

In a sample of 1481 Dutch 12-year-old twin pairs we studied the etiology of Internalizing and Externalizing problems, while taking the processes underlying agreement and disagreement between maternal and paternal ratings into account. The purpose of the analyses was to determine whether or not parental ratings of Internalizing and Externalizing behavior in 12-year-old twins conformed to rating patterns seen in younger children, to determine if there were differences in the way mothers versus fathers rated Internalizing and Externalizing behavior, and if parents rated their daughters differently than their sons.

Parent rating styles: For Externalizing behavior, both mothers and fathers rated their daughters as having fewer problems than their sons. Mother rated their daughters as having higher mean scores on Internalizing problems than their sons, where fathers reported similar rates of Internalizing behavior for both daughters and sons. Finally, the ratings for the Internalizing and Externalizing scales given by mothers were significantly higher than ratings given by fathers for both boys and girls. These reporting patterns conform with Achenbach's finding (1992).

Table 5.5.

Genetic and environmental influences, estimated using best fitting Psychometric model, for Internalizing and Externalizing Problems rated for 12-year-old twins.

	INTERNALIZING	EXTERNALIZING	
		♂	♀
Genetic			
Common	22%	48%	40%
Unique	15%	9%	11%
Shared			
Common	25%	20%	20%
Unique	11%	10%	13%
Nonshared			
Common	16%	7%	8%
Unique	11%	6%	8%

Agreement and disagreement between mother and father report. The Psychometric model fit the data best for both scales. Thus rater differences did not merely reflect measurement errors, but also indicate that parents assess different aspects of the child's behavior. These results are in accordance with previous studies (Bartels *et al.*, submitted; Hewitt *et al.*, 1992; Van der Valk *et al.*, 2001; Van der Valk *et al.*, submitted). As suggested by Achenbach *et al.* (1987) unique interactions might allow each parent to provide additional information about the child's behavior. Because no single rater may be able to provide a complete picture of the child's behavior, the implication is that it is important to collect data from multiple informants.

Variance decomposition. Genetic factors were most important for the Externalizing scale, explaining over 50% of the variance in boys and girls. Heritabilities of around 50%, without sex-differences at age 3 though, were found for 3 and 7-year-old twin pairs (Van der Valk *et al.*, 2001, Van der Valk *et al.*, submitted). Zahn-Waxler *et al.* (1996) studying 5-year-old twin pairs, Gjone et al. (1997) examining 5- to 15-year-old twin pairs, and Edelbrock *et al.* (1995) studying 7- to 15-year-old twin pairs also found that genetic influences explained about half of the variance of the Externalizing scale. At age 10 heritabilities are higher, explaining around 65% and about 50% of the variance in boys and girls separately (Bartels *et al.*, submitted). At age 12 shared environmental influences explained 25% of the variance of the Externalizing scale in boys and around 30% of the variance in girls. This again was in accordance with the shared environmental influences observed for the 3-and 7-year-old twin pairs (Van der Valk *et al.*, 2001, Van der Valk *et al.*, submitted) and the results found in the studies of Edelbrock *et al.* (1995), Gjone *et al.* (1997), and Zahn-Waxler *et al.* (1996). Apart from quantitative genetic

studies, various epidemiological studies have also demonstrated the importance of shared environmental factors in the etiology of Externalizing behaviors. Family discord and disruption, lack of affection and poor supervision all predispose to conduct problems and antisocial behavior (Rutter, 1985).

Genetic influences for the Internalizing scale accounted for 27% of the variance for 12-year-old twin pairs. In contrast, for preschool twin pairs the Internalizing scale was predominantly influenced by the child's genotype, explaining around 60% of the variance. Zahn-Waxler *et al.* (1996) also found for a sample of 5-year-old twin pairs that the genetic influences explained more than half of the variance for the Internalizing scale. It may be that the heritability for Internalizing behaviors changes with age. Shared environmental influences showed a complementary increase in influences over time, having almost no influence on the Internalizing scale of 3-year-old twin pairs (Van der Valk *et al.*, 2002) and explaining around 35% of the variance of the Internalizing scale for 7- (van der Valk *et al.*, in press), 10- (Bartels *et al.*, submitted) and 12 year-old twin pairs. A differential genetic influence for Internalizing problems of older versus younger children was also found in other studies. Gjone et al. (1997), examining a sample of twin pairs aged 5-9 and 12-15 years, found a near-significant effect of age on the genetic influence for Internalizing behaviors in terms of a decreasing genetic influence with increasing age. Also O'Connor *et al.* (1998), studying a sample of 720 siblings initially aged 10 to 18 years, found a decrease in heritability and a complementary increase in environmental influences over a three year interval for a composite score of depressive symptoms. The most plausible explanation for the differences in genetic influences on Internalizing behavior is the argument that different influences are important at different ages. At early ages, genetic influences predominate, as the child ages, and his or her environment allows for different types of behaviors, shared environmental influences become relatively more important. Contribution to the magnitude of the developmental findings are routed in cognitive psychology, developmental biology, and parenting. As brain maturation proceeds, children accrue added skills in communication, attention, and motoric behavior. These behaviors are influenced in the child's cognitive level. As Piaget (1954) has argued, children pass through successive developmental stages during which different cognitive styles are evidenced. Finally, as the child's neurodevelopment proceeds, in concert with his or her cognitive maturation, the parenting style will play an increasingly important role. Do the parents reward good behavior, set limits on negative behavior, and tolerate small failures? These positive parental attributes, as suggested by behavioral psychologists, may have different influences on a child's Internalizing behavior, than a parent who is disengaged, or disinterested, or a parent who is neglectful or abusive. In essence, as the child ages, the environment, or parent style will have a different contribution than when the child was 2 or 3 year old.

However, it may be important to realize that shared environment is not necessarily confined to the home environment. For instance, there are indications that these environmental effects are not merely shared by siblings but also by cousins (Van den Oord and Rowe, 1998; 1999). This suggests that shared environment reflects the wider community in which families are embedded as well (Bronfenbrenner, 1979; Parke and Kellam, 1994, p.3). This point has also been stressed by Harris (1995) who argues that we should think about environmental effects on development in terms of group processes where peers play an important role. That is, phenomena such as within-group assimilation and between-group contrast that increase the homogeneity of behaviors within groups and widen differences between social groups could show as shared environment in a behavior genetic analysis. Thus, the possible larger shared environmental effects in school-age versus preschool children could also reflect a developmental shift due to socialization experiences outside the home which become increasingly important as children grow older.

Sex differences in heritability were found for the Externalizing scale showing higher estimates of additive genetic effects in boys. For the Internalizing scale, girls tended to get higher scores than boys. However, no sex differences were found in our analyses of genetic and environmental estimates. For 3-year-old twin pairs (Van der Valk *et al.*, 2001) no sex differences were found, neither for the Internalizing scale nor for the Externalizing scale. For 7-year-old twin pairs sex differences were found for the Externalizing scale, parameter estimates for boys and girls were very similar though (van der Valk et al, submitted).

For the Internalizing and Externalizing scales common and unique nonshared environmental factors remained almost the same for 3-, 7-and 10-year-old twin pairs, explaining around 18% and 12% respectively of the variance of the Internalizing scale and about 11% and 7% respectively of the variance of the Externalizing scale. This indicated that parents seem to be able to rate problem behaviors of preschool children as good as problem behaviors of school-age children.

Fitting models to the observed data that explicitly incorporate rater bias and unreliability ensured that these effects could not distort estimates of the shared and nonshared environmental factors. Parameters obtained thus reflected more accurate estimates. Pure nonshared environmental influences (undistorted by error or unreliability) were found for the 12-year-old, the 10-year-old twin pairs, the 7-year-old twin pairs and for the 3-year-old twin pairs (this study; Bartels *et al.*, submitted, Van der Valk *et al.*,2001; Van der Valk *et al.*, submitted, resp.). Thus idiosyncratic experiences seem to be of importance to explain both preschool and school-age children's problem behaviors. Measurement errors and unreliability were estimated in the unique nonshared environmental factor. However, neither for the Externalizing scale nor for the Internalizing scale did this factor account for more than 11% of the variance. Rater bias was included in the estimate of the unique shared environmental

factor, accounting for at most 13% of the variance for both the Internalizing and Externalizing scale.

Conclusions. Our results confirm and extend earlier findings on the study of Internalizing and Externalizing behavior in twins. In this report of 12-year-old children, i.e., children in the peri-pubertal developmental period, we report on the differences and the similarities of the genetic and environmental influences on common childhood and adolescent behavior. Based on both mother and father reports, estimates of genetic influences on Externalizing behavior account for approximately 40-50% of the total variance. These results are consistent with the genetic influences reported by parents for earlier ages. Internalizing behavior in 12-year olds was influenced to a much smaller degree by additive genetic influences than those reported at earlier ages. We found no gender differences at age 12. Our data support the conclusion that disagreement between mother and father reports on their 12-year-old children is a results of informant specific viewpoint on the child's behavior, rather than due to measurement error.

These findings are of specific interest to the study of developmental psychopathology, because as children enter, endure, and exit puberty it will be important for parents to realize that they see and respond to different aspects of behavior. It is in fact the case, that the same child may appear different to mom and dad. More complete information on the child's behavior could also be obtained by adding teacher and self-report in studies on childhood psychopathology. Teacher reports will give insight on the child's behavior in a complete different environment than the home-environment. Besides this different environment, teacher report can be divided into two groups. One being the reports rated by teacher, who teaches both children of the twin in one classroom. The second group contains the teacher rating based on one child of the twin pair in a situation where children are taught by different teachers. Valuable information from those two groups can be obtained. Further, as the children grow older a reliable source of information on their behavior could be the self. Thus, in future studies we will extend these findings by including reports from teachers (Teacher Report Form) and the children themselves (Youth Self Report Form). By including data from multiple informants, we will be able to estimate how mothers, fathers, and teachers ratings of children compare to self-reports. By analyzing these reports on a large sample of twin data, we hope to come up with phenotyping strategies that will improve our ability to select subjects for genotyping and treatment studies.

Limitations. Our assessment instrument, the Child Behavior Checklist, does not measure DSM or ICD psychiatric diagnoses. Although, longitudinal data on the stability of Internalizing and Externalizing behavior have been reported, and the relation between Internalizing and Externalizing behavior and DSM disorders are well know, the purpose of this report is to explore the genetic and environmental influences over the course of development on these two broad syndromes, and not on DSM diagnosis. We are currently

engaged in collecting DSM data on a selected group of these twins in order to comment on diagnoses as well.

Estimates found using quantitative genetic studies do not pertain to the individual but involve average differences between individuals in the population. For other populations, or for specific individuals, other estimates may apply. This study used a nonclinical sample of Dutch twin pairs, showing problem behaviors in the normal range. Whether similar results will be obtained in clinical populations, showing more extreme problem behaviors, remains to be explored. Estimating large genetic influences in the etiology of problem behaviors, does not imply that these behaviors are not susceptible to change, for instance as a result of some interventions. The finding of genetic effects imply hereditary propensities, not predestination (Plomin and Daniels, 1986).

6

A longitudinal twin model for multiple raters: Illustrating the use of genetically informative designs for studying psychological data

---

This chapter is submitted as:

Bartels, M., Boomsma, D.I., Hudziak, J.J., Van Beijsterveldt, C.E.M., Rietveld, M.J.H., and van den Oord, E.J.C.G. (2002). A longitudinal twin model for multiple raters: Illustrating the use of genetically informative designs for studying psychological data. *Psychological Methods*, submitted.

ABSTRACT

The present article exhibits the advantages of genetically informative data. We argue that these data are not merely useful for estimating the size of genetic and environmental effects but have the potential to shed light on fundamental questions and would therefore be a useful addition to the traditional method arsenal for studying psychological data. We use longitudinal data on behavior problems in children as assessed by multiple raters and collected in a large sample of Dutch twin pairs. To analyze these data we propose a structural equation model for longitudinal and genetically informative data obtained from multiple raters. This model enables us to simultaneously investigate the etiology of developmental patterns and plausibility of different models for (dis)agreement between multiple raters. Additive genetic factors explain a large part of the stability in Internalizing and Externalizing behavior, similar assessed by both parents. Important findings are the influences of unique shared environmental influences, partly representing rater bias, on the stability in problem behaviors. Further, paternal specific view seems to be time specific only.

INTRODUCTION

One of the most important methods to study the etiology of human variation is the classical twin design. Monozygotic (MZ) twins derive from a single zygote and are therefore genetically identical. A possible way to explain differences between two members of a MZ twin pair, as indicated by less than perfect MZ twins correlations ($r_{MZ} < 1$), are environmental effects that are not shared. Examples of such nonshared environmental influences (usually indicated with symbol E) are illness, diseases, trauma, experiences at school or relationships with peers. Dizygotic (DZ) twins develop from two distinct zygotes and share on average 50% of their segregating genes, like 'ordinary' brothers and sisters. Differences between two members of a DZ twin pair can therefore result from nonshared environmental influences as well as genetic differences. A higher observed resemblance of MZ versus DZ twin pairs ($r_{MZ}>r_{DZ}$) is therefore an indication for genetic influences (A) on the trait under investigation (Martin *et al.*, 1997). The twin design also allows the study of environmental influences that are shared by members of a twin pair. Shared environmental factors (C) will create differences between families and make family members relatively more similar. Possible examples are socioeconomic level, religion, or style of parenting. Complex traits such as the ones studied in psychology are likely to be influenced by multiple genes each with a small effect. It can be shown that if these genes act in an additive manner, the DZ twin correlation is half the MZ twin correlation (Falconer and Mackay, 1996). Shared environmental influences are therefore implied for traits where $r_{DZ}>\frac{1}{2}r_{MZ}$.

Model fitting has become the standard data analytic tool in twin research (for an overview see Neale and Cardon, 1992). It basically involves solving a series of simultaneous structural equations to estimate genetic and environmental parameters that best fit observed twin correlations. This procedure also allows for the comparison of alternative models. For example, in psychopathology insight in comorbidity between certain syndromes can be obtained by fitting multivariate models. These models will then allow researchers to sort out if this comorbidity if genetically mediated, or driven by environmental factors or both.

The present article exhibits the advantages of genetically informative data. We argue that these data are not merely useful for estimating the size of genetic and environmental effects but have the potential to shed light on fundamental questions and would therefore be a useful addition to the traditional method arsenal for studying psychological data. We use longitudinal data on behavior problems in children as assessed by multiple raters and collected in a large sample of Dutch twin pairs. To analyze these data we propose a structural equation model for longitudinal and genetically informative data obtained from multiple raters. This model enables us to simultaneously investigate the etiology of

developmental patterns and plausibility of different models for (dis)agreement between multiple raters.

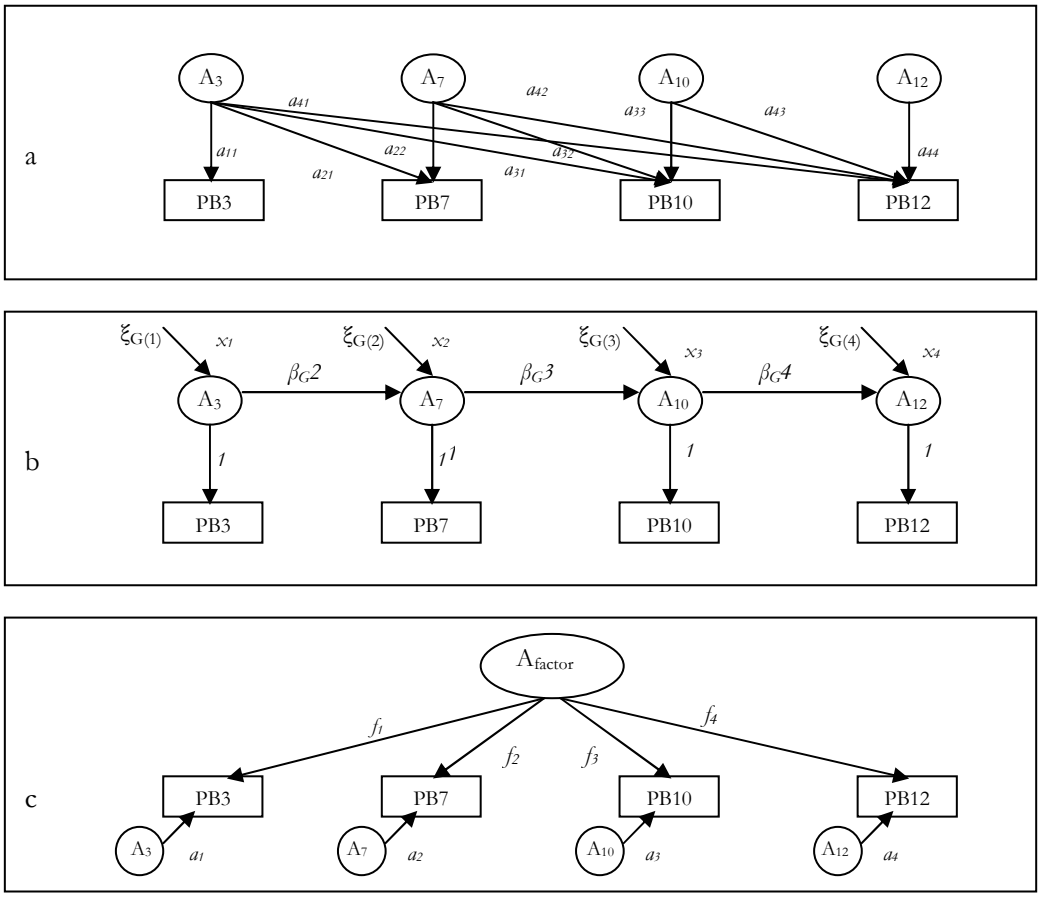
In developmental psychopathology there is considerable interest in the study of how problem behaviors develop over time. A number of articles have, for instance, reported a substantial degree of continuity (Verhulst and Van der Ende, 1992a; 1992b; Ghodsian *et al.*, 1980; Richman *et al.*, 1982; Graham and Rutter, 1973). However, because these studies have relied on phenotypic correlations, they could not distinguish whether continuity was caused by genetic or environmental influences or both. A more fundamental problem of analyzing phenotypic data is that genetic and environmental factors may display different developmental patterns. A mixture of different developmental patterns is not distinguishable at phenotypic level, so that using phenotypic data only could lead to false conclusions. An important feature of the present longitudinal twin studies is the possibility to investigate the developmental pattern of genetic and environmental factors separately.

Further, in studies on childhood psychopathology researchers commonly use parental ratings for behavioral assessment. In general parents show a reasonable agreement (e.g. interparent correlations is usually about .6; e.g. Achenbach *et al.*, 1987), which implies that parents can provide meaningful information about their child's behavior, for if parental ratings would reflect nothing but error the correlations between their ratings would be close to zero. On the other hand this interparent correlation is less than perfect. This may be explained by different forms of rater bias (a tendency of a rater to over-or underestimate behavioral problem scores consistently compared to the mean of all raters, e.g. a result of different normative standards or response tendencies) and unreliability. Another possibility is that parents are not assessing exactly the same behavior in their children. It is know that different raters can provide, each from their own perspective, somewhat different but valid and complementary information about the child's functioning (Achenbach *et al.*, 1987). Again phenotypic data do not have the statistical properties to disentangle the sources of disagreement, while using multiple raters in a genetically informative sample give rise to the opportunity to distinguish between rater bias/unreliability or unique view as the major source of disagreement. The crucial difference is that if genetic effects account for part of the rater specific variance, this unique part of the parental rating reflects the behavior of the child so that disagreement cannot be the sole result of 'error'.

In conclusion, the study of genetically related individuals provides more than estimates of genetic and environmental influences on complex traits. The statistical properties of these data yield unique information on fundamental questions in psychology such as rater (dis)agreement and developmental mechanisms not available in phenotypic designs.

Figure 6.1.

The three models used to investigate the underlying process of the development of problem behavior (PB3, PB7, PB10, PB12). 6.1a. the saturated model; 6.1b. the simplex model; 6.1c. the common factor model, with age specific influences Obviously, all three variance components (A, C, E) can be expressed in either way.



In the next section we propose a longitudinal twin model for multiple raters. We first focus on a single child and derive the contribution of different developmental mechanisms to the covariances among measurement occasions. Next, we discuss the 'Psychometric' model proposed by Hewitt (1992) for the situation where both twins of a pair are rated by both parents. Although there are alternative ways for modeling cross-sectional twin data with multiple raters, in previous studies we have found this model the most appropriate (Van der Valk *et al.*, 2001, 2002; Bartels *et al.*, submitted[a,d]). Then, we generalize this twin model to multiple measurement occasions using the developmental mechanisms discussed

in step one. Finally, we present an application of this model using longitudinal data on behavior problems in children collected in a large sample of Dutch twin pairs.

Developmental mechanisms

The three developmental mechanisms that were studied in this article are depicted in Figure 6.1. We derive a matrix $\sum$, any matrix that represents the expected contribution of each model to the covariances among measurements involving a single child. The saturated model in Figure 6.1a, also known as a Cholesky or triangular decomposition, is an unconstrained model for the (co)variances among measurement occasions. It implies the covariance structure:

$$\sum = \mathbf{X} \times \mathbf{X}' \qquad (eq.\ 6.1)$$

where $'$ indicates transposition. Matrix $\mathbf{X}$ is a $n_t \times n_t$ lower triangular matrix with n_t equal to the number of measurement occasions. For instance, for $n_t = 4$ matrix $\mathbf{X}$ would be:

$$\mathbf{X} = \begin{pmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

In the simplex model (Figure 6.1b) there are 'carry-over' or transmission effects from one measurement occasion to the subsequent one as well as effects specific to each measurement occasion. This implies the covariance structure:

$$\sum = (\mathbf{I}-\mathbf{G})^{-1} \times (\mathbf{X} \times \mathbf{X}') \times ((\mathbf{I}-\mathbf{G})^{-1})' \qquad (eq.\ 6.2)$$

Where $\mathbf{I}$ is a $n_t \times n_t$ identity matrix with elements on the main diagonal set to one. Matrix $\mathbf{G}$ is "sub" diagonal and contains the transmission effects ($\beta_G(t)$). A first order autoregressive processes is assumed so that only the elements directly below the main diagonal are estimated. For instance, with four measurement occasions $\mathbf{G}$ would equal:

$$\mathbf{G} = \begin{pmatrix} 0 & & & \\ \beta_G 2 & 0 & & \\ 0 & \beta_G 3 & 0 & \\ 0 & 0 & \beta_G 4 & 0 \end{pmatrix}$$

Matrix **X** is a $n_t \times n_t$ diagonal matrix with parameters on the main diagonal and zeroes elsewhere. It represents the new influences (ξ_{Gt}) that come into play at each measurement occasion.

$$\mathbf{X} = \begin{pmatrix} x_1 & & & \\ 0 & x_2 & & \\ 0 & 0 & x_3 & \\ 0 & 0 & 0 & x_4 \end{pmatrix}$$

In the common factor model depicted in Figure 6.1c one underlying factor with time-specific factor loadings is specified. We assume one common factor in this article but the extension to multiple factors is straightforward. To account for occasion specific variance, time specific factors are added to the model. Assuming that the common and time specific factors are uncorrelated, the expected covariance matrix equals:

$$\sum = \mathbf{Q} \times \mathbf{Q}' + \mathbf{X} \times \mathbf{X}' \qquad (eq.\ 6.3)$$

where **Q** is the $n_t \times 1$ vector with factor loadings and **X** is a $n_t \times n_t$ diagonal matrix containing the occasion-specific effects.

$$\mathbf{Q} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{pmatrix} \quad \text{and} \quad \mathbf{X} = \begin{pmatrix} a_1 & & & \\ 0 & a_2 & & \\ 0 & 0 & a_3 & \\ 0 & 0 & 0 & a_4 \end{pmatrix}$$

The saturated, simplex, and factor models result in a different pattern of covariances among measurement occasions. The saturated model is descriptive and merely estimates the variance-covariance matrix among time points. However, it is a useful model for evaluating the fit of more restricted models. For instance, if a factor model fits the data significantly poorer compared to the saturated model, it should be rejected. In a simplex model subsequent levels of the phenotype are influenced by prior levels. The implication of this autoregressive property is that effects of prior events or experiences will be larger to the extent that they happened closer in time (Guttman, 1954). The transmission model therefore predicts higher correlations among adjoining assessments than those occurring more distantly in time, the so-called simplex structure. In contrast, the factor model

assumes that the same stable factors exert their effects at each assessment and does not imply that correlations between assessments vary as a function of the length of the time lag. Thus, these models predict different patterns of longitudinal correlations, and tests can be performed to derive the underlying mechanism by comparing observed and predicted correlations.

A twin model for parental ratings

A twin model for the mother (MRT) and father (FRT) ratings is shown in Figure 6.2. The model assumes that parents partly assess the same behaviors (PT_1 and PT_2) and partly a unique aspect of their children's behaviors. The common and unique part of the ratings are influenced by additive genetic (A), shared environmental (C), and nonshared environmental (E) factors. A factor that has no subscript is a common factor affecting the ratings of both parents. Subscripts *m* and *f* are used to distinguish the factors that are specific for mother and fathers. To derive the expected covariances we first focus on a single child and write the model as a matrix equation:

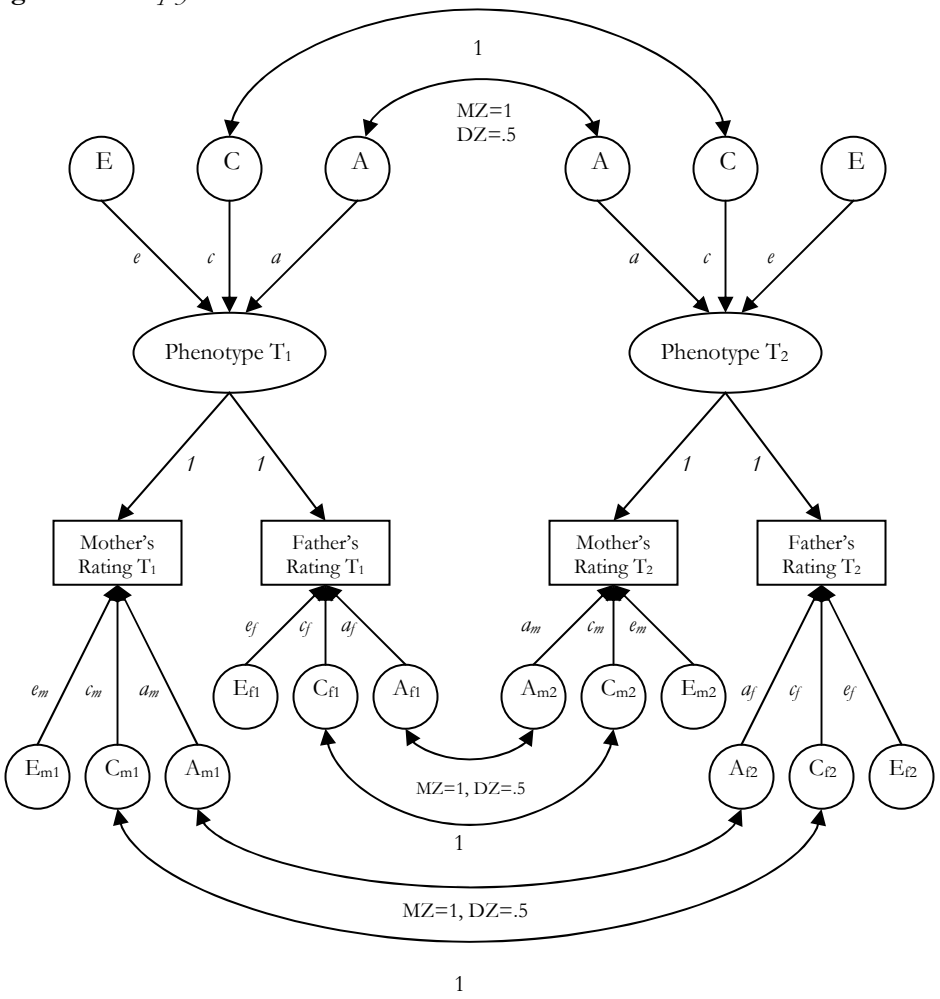
$$\begin{pmatrix} \mathrm{MRT}_1 \\ \mathrm{FRT}_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \times \begin{pmatrix} a \end{pmatrix} \times \begin{pmatrix} \mathrm{A} \end{pmatrix} + \begin{pmatrix} c \end{pmatrix} \times \begin{pmatrix} \mathrm{C} \end{pmatrix} + \begin{pmatrix} e \end{pmatrix} \times \begin{pmatrix} \mathrm{E} \end{pmatrix} + \begin{pmatrix} a_m & 0 \\ 0 & a_f \end{pmatrix} \times \begin{pmatrix} \mathrm{A_m} \\ \mathrm{A_f} \end{pmatrix} + \begin{pmatrix} c_m & 0 \\ 0 & c_f \end{pmatrix} \times \begin{pmatrix} \mathrm{C_m} \\ \mathrm{C_f} \end{pmatrix} + \begin{pmatrix} e_m & 0 \\ 0 & e_f \end{pmatrix} \times \begin{pmatrix} \mathrm{E_m} \\ \mathrm{E_f} \end{pmatrix} \qquad (eq.\ 6.4)$$

Because all genetic and environmental factors are uncorrelated, the total covariance matrix for a single child is the sum of the individual contributions. When we scale the (latent) factors by fixing their variance to one, these individual contributions can be obtained by taking the quadratic forms of the parameter matrices in equation *6.4*. This gives the following results:

$$\mathbf{A} = a^2, \ \mathbf{C} = c^2, \text{ and } \mathbf{E} = e^2$$

$$\mathbf{G} = \begin{pmatrix} a^2_m & 0 \\ 0 & a^2_f \end{pmatrix}, \ \mathbf{S} = \begin{pmatrix} \mathrm{c}^2_m & 0 \\ 0 & \mathrm{c}^2_f \end{pmatrix}, \text{ and } \mathbf{J} = \begin{pmatrix} \mathrm{e}^2_m & 0 \\ 0 & \mathrm{e}^2_f \end{pmatrix}$$

Figure 6.2. *The psychometric model*



To extent the model to a twin pair we define the data vector with parental ratings as $\mathbf{y}' = [MRT_1, FRT_1, MRT_2, FRT_2]$. By taking the expectation we obtain the expected covariance matrix for the full model in Figure 6.2:

$$\boldsymbol{\Sigma}(\mathbf{y} \times \mathbf{y}') = \mathbf{L} \times \begin{pmatrix} \mathbf{A}+\mathbf{C}+\mathbf{E} & | & r_g \otimes \mathbf{A}+\mathbf{C} \\ r_g \otimes \mathbf{A}+\mathbf{C} & | & \mathbf{A}+\mathbf{C}+\mathbf{E} \end{pmatrix} \times \mathbf{L}' + \begin{pmatrix} \mathbf{G}+\mathbf{S}+\mathbf{F} & | & r_g \otimes \mathbf{G}+\mathbf{S} \\ r_g \otimes \mathbf{G}+\mathbf{S} & | & \mathbf{G}+\mathbf{S}+\mathbf{F} \end{pmatrix}$$

(*eq. 6.5*)

where ⊗ is the Kronecker product, **L** a matrix with loadings of the common factors on the parental ratings, and r_g the correlation between the additive genetic factors of twins. The factor loading matrix is of the general form: $\mathbf{L} = \mathbf{I_t} \otimes (\mathbf{I_s} \otimes \mathbf{d})$ with $\mathbf{I_t}$ is a $n_t \times n_t$ identity matrix determined by the number of measurement occasions, $\mathbf{I_s}$ 2×2 identity matrix determined by the fact there are two children in a twin pair, and **d** is an $n_r \times 1$ vector determined by the number (n_r) of raters. Correlation r_g can be derived from quantitative genetic theory (Falconer and Mackay, 1996) and equals 1 for monozygotic twins and .5 for dizygotic twins.

Model identification

To show that the twin model is identified we express its parameters as a function of the observed statistics. We start with the common parameters. The cross-rater cross-twin covariance equals:

$$\text{Cov}(\text{MRT}_1,\text{FRT}_2) = r_g a^2 + c^2$$

Using the fact that correlation r_g equals 1 for monozygotic (MZ) and .5 for dizygotic (DZ) twins, the common genetic (a) and shared environmental (c) can be written as a function of differences in the covariances in the MZ and DZ twins groups:

$$a^2 = 2 \times (\text{Cov}(\text{MRT}_1,\text{FRT}_2)_{\text{MZ}} - \text{Cov}(\text{MRT}_1,\text{FRT}_2)_{\text{DZ}}) \qquad (eq.\ 6.6)$$

$$c^2 = 2 \times \text{Cov}(\text{MRT}_1,\text{FRT}_2)_{\text{DZ}} - \text{Cov}(\text{MRT}_1,\text{FRT}_2)_{\text{MZ}} \qquad (eq.\ 6.7)$$

Furthermore, because

$$\text{Cov}(\text{MRT}_1,\text{FRT}_1) = a^2 + c^2 + e^2$$

the nonshared environmental variance equals:

$$e^2 = \text{Cov}(\text{MRT}_1,\text{FRT}_1) - a^2 - c^2 \qquad (eq.\ 6.8)$$

To show that the rater specific effects are identified we focus on the maternal ratings but the same reasoning can be applied to the paternal ratings. We start with noting that:

$$\text{Cov}(\text{MRT}_1,\text{MRT}_2) = r_g a^2 + c^2 + r_g a^2{}_{\text{m}} + c^2{}_{\text{m}}$$

Because r_g equals 1 for monozygotic (MZ) and .5 for dizygotic (DZ) twins we can write:

$$a^2{}_{\text{m}} = 2 \times (\text{Cov}(\text{MRT}_1,\text{MRT}_2)_{\text{MZ}} - \text{Cov}(\text{MRT}_1,\text{MRT}_2)_{\text{DZ}}) - a^2 \qquad (eq.6.9)$$

$$c^2{}_{\text{m}} = 2 \times \text{Cov}(\text{MRT}_1,\text{MRT}_2)_{\text{DZ}} - \text{Cov}(\text{MRT}_1,\text{MRT}_2)_{\text{MZ}} - c^2 \qquad (eq.6.10)$$

where for a^2 and c^2 we can substitute the expressions in equations *6.6* and *6.7*. Finally, the total variance of the maternal ratings equals:

$$\text{Cov}(\text{MRT}_1,\text{MRT}_1) = a^2 + c^2 + e^2 + a^2_{\text{m}} + c^2_{\text{m}} + e^2_{\text{m}}$$

The nonshared environmental variance unique to the mother is therefore:

$$e^2_{\text{m}} = \text{Cov}(\text{MRT}_1,\text{MRT}_1) - a^2 - c^2 - e^2 - a^2_{\text{m}} - c^2_{\text{m}} \qquad (eq.\ 6.11)$$

where the terms on the right hand side of Equation *6.11* can be expressed as a function of the observed statistics by using Equations *6.6* to *6.10*.

Extending the twin model to multiple measurement occasions

In the case of multiple measurement occasions we write the data vector as $\mathbf{y}' = [\text{MRT}_{1(t=1)}, \ldots, \text{MRT}_{1(t=nt)}, \text{FRT}_{1(t=1)}, \ldots, \text{FRT}_{1(t=nt)}, \text{MRT}_{2(t=1)}, \ldots, \text{MRT}_{2(t=nt)}, \text{FRT}_{2(t=1)}, \ldots, \text{FRT}_{2(t=nt)}]$ where t indexes the measurement occasion. The expectation $E(\mathbf{y} \times \mathbf{y}')$ is again given by Equation *6.5*. However, matrices **A**, **C**, and **E** should now be replaced by one of the matrices $\sum$ in Equations *6.1* to *6.3*. This imposes a saturated, simplex, or factor structure on the covariances among the genetic and environmental factors at the measurement occasions that are common to both parents. Matrices **G**, **S**, and **J** in equation *6.5* represent the covariances among time points that involve factors unique for each rater. These unique covariances can also be modeled using saturated, simplex, or factor structures. In this case of multiple measurement occasions **G**, **J**, and **S** will have the following block diagonal structure:

$$\begin{pmatrix} \sum_{\mathbf{m}} & | & \mathbf{Z} \\ \mathbf{Z} & | & \sum_{\mathbf{f}} \end{pmatrix}$$

in which **Z** is a $n_t \times n_t$ full matrix with zeroes. Matrix $\sum$ is now subscripted *m* for mother and *f* for father to indicate that developmental mechanisms and parameter estimates may differ for both parents.

Equations similar to those presented in *6.6* to *6.11* can be used to show this model is identified. Parental ratings that apply to the same time point identify the elements on the main diagonal of $\sum$, and parental ratings that apply to different time points identify the off-diagonal element of $\sum$.

Estimation

To fit the models we used Maximum Likelihood analysis of Raw data (Lange *et al.*, 1976). RML calculates the log-likelihood for each data record. The multivariate normal density function of the observed data vector $\mathbf{y}_i$ for twin pair (i = 1..N) is:

$$g(\mathbf{y}_i; \boldsymbol{\Sigma},\boldsymbol{\mu}) = (2\pi)^{-n_{yi}/2} |\boldsymbol{\Sigma}|^{-1/2} \exp[-\tfrac{1}{2}(\mathbf{y}_i - \boldsymbol{\mu})'^t \boldsymbol{\Sigma}^{-1}(\mathbf{y}_i - \boldsymbol{\mu})]$$

where n_{yi} is the total number of ratings for twin pair i, $\boldsymbol{\mu}$ = E($\mathbf{y}$), $\boldsymbol{\Sigma}$ is given in formula *6.5*, and $|\boldsymbol{\Sigma}|$ is the determinant and $\boldsymbol{\Sigma}^{-1}$ the inverse of matrix $\boldsymbol{\Sigma}$. Because RML takes the individual data vectors as input, it can deal with missing values so that n_{yi} can range from 1 to $n_t \times n_r \times 2$. Let $\boldsymbol{\Theta}$ represent the vector of parameters used to model the means and covariances as $\boldsymbol{\mu}(\boldsymbol{\Theta})$ and $\boldsymbol{\Sigma}(\boldsymbol{\Theta})$. Parameter estimates are then obtained by maximizing the log-likelihood function given the observed data.

$$\ln L(\boldsymbol{\Theta};\mathbf{y}_i) = \sum_{i=1}^{N} \ln L_i$$

where the individual log-likelihoods equal:

$$\ln L_i = -\tfrac{1}{2}\{ n_{yi}\log(2\pi) + \log|\boldsymbol{\Sigma}(\boldsymbol{\Theta})| + (\mathbf{y}_i - \boldsymbol{\mu}(\boldsymbol{\Theta}))'\boldsymbol{\Sigma}(\boldsymbol{\Theta})^{-1}(\mathbf{y}_i - \boldsymbol{\mu}(\boldsymbol{\Theta})\}$$

The overall log-likelihood $\ln L(\boldsymbol{\Theta};\mathbf{y}_i)$ cannot be interpreted, but it can be used for model comparisons. Minus two times the difference between the log likelihoods of two nested models is chi-square distributed with the difference in estimated parameters as the degrees of freedom.

AN APPLICATION

Subjects and Measures

Longitudinal questionnaire data are collected in a large sample of Dutch twins. All participants were registered by the Netherlands Twin Registry (NTR), kept by the Department of Biological Psychology at the Vrije Universiteit in Amsterdam. Of all multiple births in the Netherlands, 40-50% is registered by the NTR (Boomsma *et al.*, 1992; Boomsma, 1998; Boomsma *et al.*, 2002). For this study, data from twins from the birth cohorts 1986 - 1993 were used. Behavioral questionnaires have been collected longitudinally at ages 3, 7, 10, and 12. Mother and father ratings were collected by making use of age-appropriate Child Behavior Checklists (CBCL 2/3, Achenbach, 1992; CBCL 4-

18, Achenbach, 1991). The CBCL is a standardized questionnaire for parents to report on the frequency of problem behavior shown by the child during the last six months. Two broadband groupings, called Internalizing and Externalizing behavior can be formed. Internalizing behavior represents withdrawn behavior and anxious/depressed behavior. Externalizing behavior represents aggressive and rule-breaking behavior. Details on the CBCL and the construction of the Internalizing and Externalizing scales can be found elsewhere (Bartels *et al.*, submitted[a, d]). At age 3 maternal ratings were available for 5,602 twin pairs, at age 7 maternal ratings were available for 5,115 twin pairs, at age 10 maternal ratings were available for 2,956 twin pairs, and at age 12 maternal ratings were available for 1,481 twin pairs. Paternal ratings were available for 3389 twin pairs at age 3, 3995 twin pairs at age 7, 2258 twin pairs at age 10, and 1155 twin pairs at age 12.

Zygosity was determined for 787 same-sex twin pairs by DNA analyses or blood group polymorphisms. For all other same-sex twin pairs zygosity was determined by discriminant analysis, using questionnaire items at each age separately. Agreement between zygosity assignment by the replies to the questions and zygosity determined by DNA markers/blood typing is around 93% (For details see Rietveld *et al.*, 2000).

Genetic modeling

The package Mx (Neale *et al.*, 1999) was used to test the different longitudinal psychometric models. A saturated model (model 1) was specified for all common and unique genetic and environmental components. This model is only descriptive and not driven by a specific hypothesis. However, it is a useful model for evaluating the fits of more restricted models. Two series of analyses were conducted. First, the significance of each variance component (unique and common) was tested. For instance, it was investigated whether disagreement between the parents was based on rater bias solely or whether real behavior, unique for each parent, was assessed. In order to make this distinction the significance of the unique additive genetic effects (A_m or A_f) was tested. If those unique genetic effects are significant, systematic effects must exist in the data that are not expected when differences parental ratings are only caused by rater bias and unreliability. It was investigated whether the disagreement between parents was based on this unique view and measurement error/unreliability solely, by testing the significance of the parental unique shared environmental influences. The significance of the influences of genetic, shared and nonshared environmental influences on the reliable trait variance (the behavior similar assessed by both parents) was tested. The only factor that was never dropped from the model was the unique nonshared environmental factor, because measurement errors are estimated in this factor.

Second, for the significant unique and common components, based on the results of model fitting series 1, continuity and the underlying developmental pattern was investigated. It was tested if the influences of a certain component is best described by time specific influences only or best represented by a longitudinal pattern (simplex or common factor). Models with time specific influences only test the significance of covariance between the ages for the variance components under investigation. A simplex model implies that there are effects specific to each age and there are 'carry-over effects' or transmission effects from one age to the subsequent age. The factor structure implies that genetic or environmental factors may exert a continuous influence from their time of onset (common factor influences). The order of model reduction and the possibilities of model specification influence the results of the parameter estimates and the goodness of fit procedure. To take this into account all variance components (A, A_m, A_f, C, C_m, C_f, or E, E_m, E_f) were analyzed separately, leaving the others expressed in a saturated model.

Table 6.1.

Multivariate model fitting for Internalizing problem behavior at age 3,7,10 and 12

INTERNALIZING							
MODEL	**-2LL**	**df**	**c.t.m.**	**χ^2**	**df**	**p**	**AIC**
1. all saturated	122775.525	50576					
2. **A_{uM}** :fixed	122875.716	50596	1	100.19	20	.00	60.19
3. **A_{uF}**: fixed	122861.215	50596	1	85.69	20	.00	45.69
4. **C_{uM}** :fixed	122900.053	50596	1	124.53	20	.00	84.53
5. **C_{uF}**: fixed	122882.667	50596	1	107.14	20	.00	67.14
6. **A_{common}** :fixed	123290.799	50596	1	515.27	20	.00	475.27
7. **C_{common}**: fixed	122862.261	50596	1	86.74	20	.00	46.74
8. **E_{common}** :fixed	124585.931	50596	1	1810.4	20	.00	1770.4

Submodels were compared by hierarchic χ^2 tests. The χ^2 statistic is computed by subtracting –2(log-likelihood) for a reduced model from that for the full model ($\chi^2 = -2LL_0 – (-2LL_1)$). In addition to the χ^2 test statistic, Akaike's Information Criterion (AIC = χ^2 – 2 x degrees of freedom) was computed. The lower the AIC the better the fit of the model to the observed data. Although the simplex model and the factor model do not

form a nested pair, they may be compared in terms of parsimony and goodness of fit because they represent alternative sets of constraints on the saturated model. Fit statistics of the reduced developmental models are compared to the saturated models. This results in a χ^2 and AIC, which are comparable for the different reduced models.

Table 6.1. -continued

Multivariate model fitting for Externalizing problem behavior at age 3,7,10 and 12

EXTERNALIZING							
MODEL	**-2LL**	**df**	**c.t.m.**	**χ^2**	**df**	**p**	**AIC**
1. all saturated	129223.468	50964					
2. **A_{uM}** :fixed	129402.159	50984	1	178.69	20	.00	138.69
3. **A_{uF}**: fixed	129289.963	50984	1	66.49	20	.00	26.49
4. **C_{uM}** :fixed	129348.411	50984	1	124.94	20	.00	84.94
5. **C_{uF}**: fixed	129437.045	50984	1	213.58	20	.00	173.58
6. **A_{common}** :fixed	130355.095	50984	1	1131.63	20	.00	1091.6
7. **C_{common}**: fixed	129368.438	50984	1	144.97	20	.00	104.97
8. **E_{common}** :fixed	131076.650	50984	1	1853.18	20	.00	1813.1

RESULTS

The saturated, Cholesky Decomposition, model without restrictions (model 1) was taken as a reference for evaluating changes in χ^2 and associated degrees of freedom of more parsimonious models. Because of the fact that the order of model reductions could influence the results of the goodness of fit procedure all variance components were analyzed separately.

Table 6.1 represents the model fitting results for testing the significance of all common and unique genetic and environmental components for both Internalizing and Externalizing behavior. All variance components are significant, indicated by the poorer fit of the reduced models. The χ^2 changes dramatically after fixing one of the variance components to zero and high AIC values are found. Thus, each parent provides specific information on the behavior of his or her child (model 2 and 3). Further, significant unique shared environmental influences seem to exist, partly representing rater bias

(model 4 and 5). Finally, significant additive genetic, shared environmental and nonshared environmental influences on the reliable trait variance are found (model 6, 7, and 8).

The percentages of the total age specific variance and the total between age covariances decomposed in common and unique additive genetic, shared environmental, and nonshared environmental factors based on the saturated model are presented in Table 6.2. Common factors, influencing behaviors similarly assessed by both parents, were more important than unique factors, influencing behaviors uniquely assessed by one parent. Common additive genetic factors are most important in explaining individual differences in both Internalizing and Externalizing behavior at ages 3, 7, 10, and 12. However, for Internalizing behavior a decrease in common additive genetic influences is found over the years. A complementary increase in common shared environmental influences is found. These results are comparable to previous studies conducted with overlapping samples (Van der Valk *et al.*, 1998; 2001; 2002; Bartels *et al.*, submitted[a, d, e]). Variance and covariance decomposition for boys and girls separately are presented in Appendix 6.1.

Important in Table 6.2 (off diagonal) are the influences of common and unique genetic and environmental factors on the covariances. For Internalizing behavior, based on mother ratings, stability, represented in these covariances, can be explained both by common additive genetic and common shared environmental influences. Common nonshared environmental influences seem of less important for stability in problem behavior, represented by very low influences on the covariances. Based on the mother ratings stability of Internalizing behavior over the years is for 41% ((50% + 45% + 45% +39% + 34% + 31%) /6) explained by common additive genetic influences, on average. 28% of this stability can be explained by common shared environmental factors.

A salient finding is the significant and rather high influence of unique shared environmental influences on the variances as well as the covariances. This influence can represent two components. First, it can represent 'real' shared environmental influences, uniquely assessed by one of the parents. Second, it can represent rater bias. The fact that a relative large part of the total variance and covariance is explained by this unique shared environmental factor point more into the direction of rater bias. However, some parental specific shared environmental influence is expected. About 19% of the covariance in Internalizing behavior, based on the mother ratings, is accounted for by these unique shared environmental influences. So, a significant part of the stability is accounted for by rater bias. For Internalizing behavior based on father ratings a similar picture emerges (see Table 6.2).

Table 6.2.

Percentages of the total variances (diagonal) and covariances (off diagonal) for Internalizing problem behavior in boys and girls based mother and father ratings separately using the best fitting model.

	INTERNALIZING									
	Mother					**Father**				
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
A_c[a]	**3**	.48				**3**	.49			
	7	.50	.33			**7**	.59	.35		
	10	.45	.39	.26		**10**	.56	.47	.30	
	12	.45	.34	.31	.25	**12**	.54	.39	.36	.27
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
C_c[b]	**3**	.06				**3**	.06			
	7	.24	.13			**7**	.28	.14		
	10	.29	.22	.19		**10**	.36	.26	.21	
	12	.35	.29	.27	.20	**12**	.42	.33	.31	.22
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
E_c[c]	**3**	.14				**3**	.14			
	7	.03	.15			**7**	.04	.16		
	10	.01	.11	.15		**10**	.00	.14	.17	
	12	.00	.12	.16	.14	**12**	.00	.13	.18	.15
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
A_u[d]	**3**	.13				**3**	.09			
	7	.02	.12			**7**	.08	.12		
	10	.02	.09	.11		**10**	.09	.00	.07	
	12	.00	.04	.09	.11	**12**	.00	.00	.02	.00
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
C_u[e]	**3**	.05				**3**	.09			
	7	.19	.15			**7**	.02	.12		
	10	.24	.17	.14		**10**	.06	.15	.15	
	12	.24	.16	.12	.17	**12**	.18	.15	.13	.08
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
E_u[f]	**3**	.14				**3**	.12			
	7	.02	.12			**7**	.00	.11		
	10	.00	.03	.15		**10**	.00	.01	.10	
	12	.05	.03	.06	.12	**12**	.01	.02	.00	.11

[a] Additive genetic influence on the reliable trait variance; [b] shared environmental influence on the reliable trait variance; [c] nonshared environmental influence on the reliable trait variance; [d] parental unique genetic influences; [e] parental unique shared environmental variance; [f] parental unique nonshared environmental variance.

Table 6.2. - continued

Percentages of the total variances (diagonal) and covariances (off diagonal) for Externalizing problem behavior in boys and girls based mother and father ratings separately using the best fitting model.

	EXTERNALIZING									
	Mother					Father				
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
$\mathbf{A_c}$[a]	**3**	.44				**3**	.45			
	7	.53	.47			**7**	.62	.49		
	10	.49	.53	.48		**10**	.54	.58	.50	
	12	.48	.56	.58	.45	**12**	.56	.64	.64	.48
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
$\mathbf{C_c}$[b]	**3**	.20				**3**	.21			
	7	.23	.16			**7**	.27	.17		
	10	.28	.21	.16		**10**	.31	.23	.16	
	12	.28	.19	.16	.18	**12**	.33	.22	.18	.19
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
$\mathbf{E_c}$[c]	**3**	.10				**3**	.11			
	7	.04	.09			**7**	.05	.09		
	10	.03	.08	.09		**10**	.03	.08	.10	
	12	.03	.05	.07	.07	**12**	.03	.06	.08	.08
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
$\mathbf{A_u}$[d]	**3**	.09				**3**	.04			
	7	.06	.12			**7**	.00	.07		
	10	.09	.12	.12		**10**	.00	.03	.04	
	12	.08	.09	.10	.15	**12**	.00	.00	.00	.00
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
$\mathbf{C_u}$[e]	**3**	.09				**3**	.12			
	7	.12	.10			**7**	.05	.12		
	10	.11	.05	.09		**10**	.13	.08	.15	
	12	.12	.08	.06	.08	**12**	.12	.09	.14	.14
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
$\mathbf{E_u}$[f]	**3**	.07				**3**	.08			
	7	.01	.06			**7**	.01	.06		
	10	.00	.01	.07		**10**	.00	.00	.06	
	12	.00	.02	.02	.07	**12**	.01	.00	.01	.06

[a] Additive genetic influence on the reliable trait variance; [b] shared environmental influence on the reliable trait variance; [c] nonshared environmental influence on the reliable trait variance; [d] parental unique genetic influences; [e] parental unique shared environmental variance; [f] parental unique nonshared environmental variance.

For Externalizing behavior common additive genetic factors and common shared environmental factors also seem to be the main source of stability. For Externalizing behavior, based on mother ratings, 53% of the stability is explained by common additive genetic factors, on average. Common shared environmental factors explain 23% of the total covariance on average. For Externalizing behavior also, a significant part of the covariances is accounted for by unique shared environmental factors, explaining 9% of the covariances, on average. For Externalizing behavior based on father a similar picture emerges (see Table 6.2.). Common nonshared environmental factors, unique additive genetic factors and unique nonshared environmental factors are less important for stability in Internalizing and Externalizing behavior, represented by the low estimates in Table 6.2.

Developmental patterns

Table 6.3 represents the results of the fit of time specific and distinct longitudinal models. By using a model with time specific influences only it is tested whether the various variance components account for stability in problem behavior. If it is found that the influence of a certain variance component can be expressed by time specific influences, only change in problem behavior throughout development is expected as a result of the influence of this particular variance component.

The best fitting model for each component is boldfaced. Stability in 'reliable assessed' problem behavior, represented by behavior similar assessed by both parents, seem to be accounted for by additive genetic, shared environmental and nonshared environmental influences. Further, the maternal specific view is important for stability, represented by unique additive genetic and unique shared environmental influences. Part of these unique shared environmental influences will represent 'real' behavior, while part of this influence will be accounted for by rater bias. The unique aspects of behavior solely observed by the father seem to be time specific and not important for stability in problem behavior. For father ratings significant influence on the covariance is observed for unique shared environmental influences only. Because of the fact that no continuity is observed in the paternal unique additive genetic influences it is expected that the paternal unique shared environmental influences mainly represent rater bias.

For the factors that show continuity, represented by a poorer fit of a model with time specific factors only, it is tested whether a simplex model or a common factor describes the developmental process best. For the common factors A, C and E, representing Externalizing behavior similar assessed by both parents, the developmental pattern is best represented by a simplex model (model 3, 6, 9).

Table 6.3.

Multivariate model fitting for Internalizing problem behavior at age 3,7,10 and 12

INTERNALIZING							
MODEL	**-2LL**	**df**	**c.t.m.**[a]	**χ^2**	**df**	**p**	**AIC**
1. all saturated model	122775.525	50576					
2. **A**: Time Specific	122976.893	50588	1	201.37	12	.00	177.37
3. **A**: Simplex Structure	**122780.547**	**50582**	**1**	**5.02**	**6**	**.45**	**-6.98**
4. **A**: Factor Structure	122777.723	50580	1	2.19	4	.70	-5.80
5. **C**: Time Specific	122848.048	50588	1	72.52	12	.00	48.52
6. **C**: Simplex Structure	**122780.435**	**50582**	**1**	**4.91**	**6**	**.56**	**-7.09**
7. **C**: Factor Structure	122776.504	50580	1	.979	4	.91	-7.02
8. **E**: Time Specific	123076.546	50588	1	301.02	12	.00	277.02
9. **E**: Simplex Structure	**122781.993**	**50582**	**1**	**6.47**	**6**	**.37**	**-5.53**
10. **E**: Factor Structure	122780.561	50580	1	5.04	4	.28	-2.96
11. $\mathbf{A_m}$: Time Specific	122793.968	50588	1	18.44	12	.10	-5.56
12. $\mathbf{A_m}$: Simplex Structure	**122779.991**	**50582**	**1**	**4.47**	**6**	**.61**	**-7.53**
13. $\mathbf{A_m}$: Factor Structure	122778.109	50580	1	2.58	4	.63	-5.42
14. $\mathbf{A_f}$: Time Specific	**122781.740**	**50588**	**1**	**6.22**	**12**	**.91**	**-17.78**
15. $\mathbf{A_f}$: Simplex Structure	122794.425	50582	1	18.90	6	.00	6.90
16. $\mathbf{A_f}$: Factor Structure	122776.437	50580	1	.91	4	.92	-7.09
17. $\mathbf{C_m}$: Time Specific	122842.775	50588	1	67.25	12	.00	43.25
18. $\mathbf{C_m}$: Simplex Structure	122791.595	50582	1	16.07	6	.01	4.07
19. $\mathbf{C_m}$: Factor Structure	**122778.289**	**50580**	**1**	**2.76**	**4**	**.60**	**-5.24**
20. $\mathbf{C_f}$: Time Specific	122818.176	50588	1	42.65	12	.00	18.65
21. $\mathbf{C_f}$: Simplex Structure	122782.192	50582	1	6.667	6	.35	5.33
22. $\mathbf{C_f}$: Factor Structure	**122776.659**	**50580**	**1**	**1.134**	**4**	**.89**	**-6.87**
23. $\mathbf{E_m}$: Time Specific	122829.756	50588	1	54.23	12	.00	30.23
24. $\mathbf{E_m}$: Simplex Structure	122801.064	50582	1	25.54	6	.00	13.54
25. $\mathbf{E_m}$: Factor Structure	122960.449	50580	1	184.92	4	.00	176.92
26. $\mathbf{E_f}$: Time Specific	**122794.425**	**50588**	**1**	**18.90**	**12**	**.09**	**-5.10**
27. $\mathbf{E_f}$: Simplex Structure	122785.044	50582	1	9.52	6	.15	-2.48
28. $\mathbf{E_f}$: Factor Structure	122790.397	50580	1	14.87	4	.05	6.87

[a] c.t.m. = compared to model

Table 6.3. - continued

Multivariate model fitting for Externalizing problem behavior at age 3,7,10 and 12

EXTERNALIZING							
MODEL	**-2LL**	**df**	**c.t.m.**[a]	**χ^2**	**df**	**p**	**AIC**
1. all saturated model	129223.468	50964					
2. **A**: Time Specific	129871.175	50976	1	647.71	12	.00	623.71
3. **A**: Simplex Structure	**129226.374**	**50970**	**1**	**2.91**	**6**	**.82**	**-9.09**
4. **A**: Factor Structure	129241.623	50968	1	18.16	4	.00	10.15
5. **C**: Time Specific	129302.599	50976	1	79.13	12	.00	55.13
6. **C**: Simplex Structure	**129229.710**	**50970**	**1**	**6.24**	**6**	**.40**	**-5.76**
7. **C**: Factor Structure	129228.502	50968	1	5.03	4	.28	-2.97
8. **E**: Time Specific	129564.838	50976		341.37	12	.00	317.37
9. **E**: Simplex Structure	**129226.442**	**50970**	**1**	**2.97**	**6**	**.81**	**-9.03**
10. **E**: Factor Structure	129231.807	50968	1	8.34	4	.08	.34
11. $\mathbf{A_m}$: Time Specific	129282.331	50976	1	58.86	12	.00	34.86
12. $\mathbf{A_m}$: Simplex Structure	129237.262	50970	1	13.79	6	.03	1.79
13. $\mathbf{A_m}$: Factor Structure	**129230.039**	**50968**	**1**	**6.57**	**4**	**.16**	**-1.43**
14. $\mathbf{A_f}$: Time Specific	**129237.775**	**50976**	**1**	**14.31**	**12**	**.28**	**-9.69**
15. $\mathbf{A_f}$: Simplex Structure	129240.775	50970	1	17.31	6	.01	5.31
16. $\mathbf{A_f}$: Factor Structure	129224.851	50968	1	1.38	4	.85	-6.62
17. $\mathbf{C_m}$: Time Specific	129271.107	50976	1	47.64	12	.00	23.64
18. $\mathbf{C_m}$: Simplex Structure	129236.495	50970	1	13.03	6	.04	1.03
19. $\mathbf{C_m}$: Factor Structure	**129226.141**	**50968**	**1**	**2.67**	**4**	**.61**	**-5.33**
20. $\mathbf{C_f}$: Time Specific	129261.961	50976	1	38.49	12	.00	14.49
21. $\mathbf{C_f}$: Simplex Structure	129243.029	50970	1	19.56	6	.00	7.56
22. $\mathbf{C_f}$: Factor Structure	**129225.626**	**50968**	**1**	**2.16**	**4**	**.71**	**-5.84**
23. $\mathbf{E_m}$: Time Specific	129266.504	50976	1	43.04	12	.00	19.04
24. $\mathbf{E_m}$: Simplex Structure	**129228.106**	**50970**	**1**	**4.64**	**6**	**.59**	**-7.36**
25. $\mathbf{E_m}$: Factor Structure	129388.835	50968	1	165.37	4	.00	157.37
26. $\mathbf{E_f}$: Time Specific	129240.775	50976	1	17.31	12	.14	-6.69
27. $\mathbf{E_f}$: Simplex Structure	**129228.106**	**50970**	**1**	**4.64**	**6**	**.59**	**-7.36**
28. $\mathbf{E_f}$: Factor Structure	129249.248	50968	1	25.78	4	.00	17.78

[a] c.t.m. = compared to model

For Internalizing behavior, similar assessed by both parents, the picture is less clear. As can be seen in Table 6.3, for common additive genetic and common shared environmental influences on Internalizing behavior the factor (model 7) and the simplex (model 6) model gave almost identical fits to the data. It depends on the fit index which developmental pattern is considered to be best. Based on the AIC, the developmental pattern for the common factors A, C, and E, representing Internalizing behavior similar assessed by both parents, is best represented by a simplex model (model 3, 6, 9). Based on the p-value, a factor model for both additive genetic and shared environmental influences is preferred (model 4 and 7). The simplex model, though, is a more parsimonious model and therefore preferred over a factor model. This parsimony is reflected in a lower AIC for the simplex model than for the factor model, suggesting a better fit of the simplex model. It should be noted, however, that it is an arbitrary choice, based on the preferred fit index.

For the parental unique views on behavior the distinct components show different developmental patterns. From the paternal unique view, only shared environmental influences are of significance for stability. The developmental pattern is best described by a factor structure (model 22). Because of the fact that this is the only paternal specific influence on continuity it is expected that it mainly represents rater bias. The maternal shared environmental influences, on both problem behaviors, are also best described by a factor structure. This suggests that a large part of this influences is rater bias. For the maternal view the significant additive genetic influences on stability in Externalizing behavior a factor model gave the best fit. For the maternal unique nonshared environmental influences the developmental pattern in Internalizing behavior is best described by a saturated model while this influence on Externalizing behavior is best described by a simplex structure.

DISCUSSION

In this article we argue that genetically informative data are not merely useful for estimating the size of genetic and environmental effects but have the potential to shed light on more fundamental questions and would be a useful addition to the traditional method arsenal for studying psychological data. To illustrate we propose a structural equation model for longitudinal and genetically informative data obtained from multiple raters and fit the model to a large sample of Dutch twin pairs. This models enabled us to investigate the etiology of developmental patterns and plausibility of different models for (dis)agreement between multiple raters.

Our model allowed us to decompose the variance into common additive genetic, shared environmental and nonshared environmental influences and unique additive genetic, shared environmental and nonshared environmental influences. The 'common'

components influence behavior similar assessed by both parents, i.e. reliable trait variance. The 'unique' components represent influences on behavior uniquely assessed by one of the parents. Based on the interparent correlation of about .6, the common factors were more important than unique factors in explaining individual differences in problem behavior. The variance of problem behavior at the distinct ages is mainly accounted for by common additive genetic influences. Based on the mother ratings 48%, 33%, 26%, and 25% of the total variance in Internalizing behavior is accounted for by additive genetic factors at ages 3, 7, 10, and 12 respectively. For Externalizing behavior based on the mother ratings 44%, 47%, 48%, and 45% of the total variance is accounted for by additive genetic factors at ages 3, 7, 10, and 12 respectively. For Internalizing behavior a decrease in additive genetic influences is found over the years. A complementary increase in common shared environmental influences is found.

This study shows that stability is mainly accounted for by common additive genetic and shared environmental influences. Stability in Internalizing behavior, based on mother ratings is for 41% accounted for by common additive genetic factors and for 28% by common shared environmental influences. In Externalizing behavior, based on mother ratings, stability is for 53% accounted for by common additive genetic factors, while 23% is accounted for by common shared environmental influences. Although there were exceptions the developmental process for the common additive genetic, shared environmental and nonshared environmental influences seemed in most cases best described by a simplex structure. This simplex-like continuity for genetic and environmental influences assumes that successive levels of functioning were causally linked and that earlier experiences and/or genetic effects affected later maladjustment.

Another finding is that the maternal and paternal unique shared environmental influences are best described by a factor structure. This structure as well as the fact that compared to the common shared environmental effects, the unique effects are relatively large seems to suggest a significant influence of rater bias on stability in problem behaviors. In interpreting results of longitudinal studies it should therefore be realized that part of the stability is caused by stability in the 'rater bias' instead of stability in the 'real' behavior of the child.

The significant influences of additive genetic factors on the reliable trait variance indicate a possible innate vulnerability to childhood psychopathology. The influences of common nonshared environmental influences suggest the importance of pure idiosyncratic experiences. Significant influences of environment shared by both members of a twin pair are represented by the common shared environmental factor. The significant influence of unique additive genetic influences represent the fact that parents assess unique aspects of their child's behavior and so the use of multiple raters adds rater specific

information. Unique shared environmental influences partly represent 'real' shared environmental influences on behavior uniquely assessed by one parent, but also represents rater bias. Measurement error is mainly accounted for by unique nonshared environmental influences.

The general model can be readily extended to more than two raters, for instance teacher ratings can be added to the model. Information from teachers can shed a different light on rater bias and specific view. First teachers do see a lot of children of the same age, so they might be more capable to judge a child's behavior as being relatively normal or deviant. Further, they interact with the child in a school environment that could result in a different view on behavior in comparison to behavior art home, rated by parents.

Phenotypic studies on change and stability in problem behavior also indicate Internalizing and Externalizing behavior to be highly stable. Verhulst and Van der Ende (1992a, 1992b) studied stabilities in problem behavior in a sample of 936 Dutch 4- to 11-year old children. This study is especially interesting because children were almost the same age range and from the same Dutch population as the children in the present article. The average observed stability coefficients for the two-four-, and six-year time intervals were, respectively, .53, .48, .42. The decrease of the coefficients with increasing time interval suggests a simplex pattern as the underlying developmental process. However, the stability between the distinct ages is higher than would be expected based on a simplex structure solely. This same pattern of stability coefficient is also found in a national sample of 16,000 British children at ages 7, 11, and 16 years (Ghodsian *et al.*, 1980).

The present study sheds new light on these findings in several ways. First, it shows that stability is mainly accounted for by common additive genetic factors, common shared environmental factors, and unique shared environment, partly representing rater bias. This knowledge about the causes of stability might be useful for prevention. For instance, suppose that the results found in the present study can be generalized and that especially children with a high genetic liability or children who continue to experience adverse shared environments are at risk for later maladjustment. For these children, a 'wait and see' policy would be inappropriate and an active intervention would be required. On the other hand, for behavior problems that mainly have nonshared environmental etiologies, intervention would be superfluous because these problem would be of transient nature.

Second, nine to 20% of the stability in Internalizing and Externalizing problem behavior is accounted for by unique shared environmental influences, partly representing rater bias, displaying a factor structure. In other words, the developmental pattern of rater bias (a common factor structure) could have resulted in misleading results when the influences of 'real' shared environment and rater bias cannot be disentangled. This may partly explain the finding of the relative higher stability coefficients in phenotypic studies

than would be expected on a simplex structure for development. So, using genetically related individuals and multiple raters give rise to the capability of distinguishing the distinct sources influencing stability in problem behavior.

In general it is recommended to use multiple raters (see also Achenbach *et al.*, 1987), because different raters are assumed to provide, each from their own perspective, unique and complementary information about the children's functioning. Using more than one rating will give more reliable results by decreasing measurement error. However, when studying developmental patterns to predict and prevent maladjustment the use of father ratings is questionable. Results of this study indicate that the influences of the paternal specific view (A_f) are time specific only. Fathers seem to attribute additional information on the child's behavior, however no continuity in this view is observed. It could be the case that these age specific paternal views are a representation of trivial behavioral fluctuations rather than real father-child specific interaction. Further, mothers specific view show continuity over the years, the influence on the covariance, though, is very low. Thus, the use of multiple raters is necessary to correct for the influence of rater bias. In the scope of prediction and prevention father raters are less valuable.

The present article shows one example of the advantages of genetically informative data for analyzing psychological data. Other applications are conceivable. For instance, a commonly used approach to study the structure of psychological tests is factor analysis. Its equivalent in a genetically informative design is a model where the genetic and environmental effects influence the latent factor. An implicit assumption of this model, and consequently of phenotypic factor analyses, is that genetic and environmental factors cause similar patterns of item correlations and that the relative importance of genetic and environmental factors is identical for all items. Assume, for instance, a three item test with loadings λ_1, λ_2, and λ_3, that the genetic effect on the factor is a, and the shared environmental effect is c. Because the genetic and shared environmental factors affect the items via the "phenotypic" factor, their effects on the items equal $a\lambda_1$, $a\lambda_2$, and $a\lambda_3$ and $c\lambda_1$, $c\lambda_2$, and $c\lambda_3$ respectively. Except for the proportion $(a/c)^2$, the genetic and environment effects on the items and their inter correlations are therefore identical. However, in reality this model may be wrong because genetic and environmental factors affect the items and different ways. For instance, in the context for neuroticism (Tyrer, 1985), it has been speculated that genetic factors mainly determine a general vulnerability and that the environment codes for the specific symptoms. In this case, only genetic factors would contribute to the correlations between the symptoms and the relative importance of genetic and environmental factors is likely to vary across symptoms. Regardless of the validity of this specific theoretical model, there is empirical support for the notion that genetic and environmental factor vary across symptoms of neuroticisms (Martin *et al.*,

1979). With only phenotypic data it will very difficult to recover the correct factor structure if genetic and environmental factors affect items via different pathways. We speculate that in this situation the factor structure will be unclear and that the phenotypic solution will tend to the extract the genetic or environmental factor that explains most of the variance. A scale derived on the basis of this factor solution will therefore mainly tap this component and systematically ignore the other sources of variation. By using genetically informative data it will be possible to derive the correct factor structure. Further, factor estimation procedure could be used to derive genetic and environmental test scores (Boomsma *et al.*, 1990). Such a differential scoring at the level of etiology might be useful for several reasons. For instance, it could be used to obtain more refined diagnoses of children's problem behaviors which could be important to optimize response to treatment.

Appendix 6.1.

Percentages of the total variances (diagonal) and covariances (off diagonal) for Internalizing problem behavior in boys and girls based mother and father ratings separately using the best fitting model.

	INTERNALIZING in BOYS									
	Mother					**Father**				
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
A_c[a]	**3**	.51				**3**	.53			
	7	.56	.36			**7**	.71	.38		
	10	.48	.44	.28		**10**	.71	.53	.32	
	12	.49	.39	.35	.28	**12**	.66	.44	.39	.31
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
C_c[b]	**3**	.03				**3**	.04			
	7	.16	.11			**7**	.21	.11		
	10	.23	.20	.18		**10**	.33	.23	.21	
	12	.26	.26	.25	.18	**12**	.36	.29	.28	.20
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
E_c[c]	**3**	.15				**3**	.16			
	7	.03	.15			**7**	.04	.16		
	10	.00	.10	.16		**10**	.00	.12	.18	
	12	.00	.12	.16	.14	**12**	.00	.14	.18	.16
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
A_u[d]	**3**	.09				**3**	.11			
	7	.03	.13			**7**	.05	.10		
	10	.04	.13	.12		**10**	.00	.00	.05	
	12	.00	.12	.10	.13	**12**	.00	.00	.01	.00
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
C_u[e]	**3**	.08				**3**	.05			
	7	.20	.13			**7**	.01	.12		
	10	.27	.13	.12		**10**	.12	.12	.13	
	12	.31	.10	.10	.16	**12**	.13	.13	.12	.07
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
E_u[f]	**3**	.13				**3**	.12			
	7	.01	.12			**7**	.00	.13		
	10	.00	.00	.14		**10**	.00	.04	.11	
	12	.06	.01	.04	.10	**12**	.00	.03	.02	.10

[a] Additive genetic influence on the reliable trait variance; [b] shared environmental influence on the reliable trait variance; [c] nonshared environmental influence on the reliable trait variance; [d] parental unique genetic influences; [e] parental unique shared environmental variance; [f] parental unique nonshared environmental variance.

Appendix 6.1. - continued

Percentages of the total variances (diagonal) and covariances (off diagonal) for Internalizing problem behavior in boys and girls based mother and father ratings separately using the best fitting model.

	INTERNALIZING in GIRLS									
	Mother					**Father**				
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
A_c[a]	**3**	.45				**3**	.46			
	7	.43	.29			**7**	.47	.31		
	10	.43	.35	.24		**10**	.48	.42	.27	
	12	.41	.27	.28	.21	**12**	.44	.32	.34	.23
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
C_c[b]	**3**	.09				**3**	.09			
	7	.31	.16			**7**	.34	.17		
	10	.33	.22	.19		**10**	.37	.27	.22	
	12	.43	.33	.27	.23	**12**	.46	.39	.33	.25
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
E_c[c]	**3**	.13				**3**	.13			
	7	.03	.15			**7**	.04	.16		
	10	.00	.12	.14		**10**	.01	.15	.16	
	12	.00	.11	.15	.14	**12**	.00	.13	.19	.15
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
A_u[d]	**3**	.14				**3**	.01			
	7	.04	.11			**7**	.00	.11		
	10	.00	.06	.10		**10**	.00	.00	.06	
	12	.00	.00	.06	.06	**12**	.00	.00	.02	.13
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
C_u[e]	**3**	.04				**3**	.19			
	7	.17	.16			**7**	.15	.15		
	10	.22	.20	.17		**10**	.22	.23	.19	
	12	.26	.19	.16	.21	**12**	.12	.17	.16	.13
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
E_u[f]	**3**	.14				**3**	.11			
	7	.02	.12			**7**	.03	.10		
	10	.01	.05	.16		**10**	.00	.00	.09	
	12	.05	.09	.08	.14	**12**	.05	.01	.00	.11

[a] Additive genetic influence on the reliable trait variance; [b] shared environmental influence on the reliable trait variance; [c] nonshared environmental influence on the reliable trait variance; [d] parental unique genetic influences; [e] parental unique shared environmental variance; [f] parental unique nonshared environmental variance.

Appendix 6.1. - continued

Percentages of the total variances (diagonal) and covariances (off diagonal) for Externalizing problem behavior in boys and girls based mother and father ratings separately using the best fitting model.

	EXTERNALIZING in BOYS									
	Mother					**Father**				
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
A_c[a]	**3**	.48				**3**	.48			
	7	.58	.49			**7**	.66	.50		
	10	.55	.55	.52		**10**	.60	.61	.53	
	12	.51	.58	.61	.47	**12**	.59	.67	.66	.49
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
C_c[b]	**3**	.18				**3**	.18			
	7	.20	.15			**7**	.23	.16		
	10	.25	.20	.14		**10**	.27	.22	.15	
	12	.25	.17	.16	.19	**12**	.29	.19	.17	.20
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
E_c[c]	**3**	.10				**3**	.10			
	7	.04	.09			**7**	.04	.09		
	10	.02	.06	.08		**10**	.02	.07	.09	
	12	.02	.05	.06	.06	**12**	.02	.06	.07	.07
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
A_u[d]	**3**	.05				**3**	.05			
	7	.04	.11			**7**	.03	.09		
	10	.06	.12	.12		**10**	.00	.03	.02	
	12	.07	.13	.08	.13	**12**	.00	.02	.00	.02
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
C_u[e]	**3**	.13				**3**	.11			
	7	.14	.09			**7**	.03	.11		
	10	.13	.04	.07		**10**	.13	.06	.16	
	12	.13	.06	.06	.09	**12**	.10	.06	.14	.14
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
E_u[f]	**3**	.07				**3**	.08			
	7	.00	.06			**7**	.02	.06		
	10	.00	.01	.06		**10**	.00	.00	.05	
	12	.02	.02	.03	.06	**12**	.03	.00	.00	.04

[a] Additive genetic influence on the reliable trait variance; [b] shared environmental influence on the reliable trait variance; [c] nonshared environmental influence on the reliable trait variance; [d] parental unique genetic influences; [e] parental unique shared environmental variance; [f] parental unique nonshared environmental variance.

Appendix 6.1. - continued

Percentages of the total variances (diagonal) and covariances (off diagonal) for Externalizing problem behavior in boys and girls based mother and father ratings separately using the best fitting model.

	EXTERNALIZING in GIRLS									
	Mother					**Father**				
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
A_c[a]	**3**	.41				**3**	.42			
	7	.48	.46			**7**	.57	.48		
	10	.41	.48	.41		**10**	.45	.51	.43	
	12	.41	.52	.52	.40	**12**	.50	.58	.57	.43
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
C_c[b]	**3**	.22				**3**	.23			
	7	.27	.17			**7**	.31	.18		
	10	.34	.25	.19		**10**	.38	.27	.20	
	12	.34	.24	.21	.21	**12**	.41	.27	.23	.22
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
E_c[c]	**3**	.11				**3**	.11			
	7	.05	.09			**7**	.06	.09		
	10	.05	.09	.10		**10**	.06	.10	.11	
	12	.03	.06	.08	.09	**12**	.04	.07	.08	.09
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
A_u[d]	**3**	.11				**3**	.04			
	7	.08	.13			**7**	.00	.03		
	10	.13	.09	.09		**10**	.00	.05	.05	
	12	.08	.06	.11	.11	**12**	.00	.00	.00	.02
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
C_u[e]	**3**	.08				**3**	.13			
	7	.10	.10			**7**	.08	.15		
	10	.08	.08	.13		**10**	.14	.08	.14	
	12	.13	.09	.08	.13	**12**	.08	.10	.12	.15
		3	**7**	**10**	**12**		**3**	**7**	**10**	**12**
E_u[f]	**3**	.07				**3**	.07			
	7	.02	.06			**7**	.00	.07		
	10	.00	.01	.08		**10**	.00	.00	.07	
	12	.00	.03	.02	.07	**12**	.00	.00	.03	.08

[a] Additive genetic influence on the reliable trait variance; [b] shared environmental influence on the reliable trait variance; [c] nonshared environmental influence on the reliable trait variance; [d] parental unique genetic influences; [e] parental unique shared environmental variance; [f] parental unique nonshared environmental variance.

7

Genetic and Environmental Influences on the Development of Intelligence

This chapter is published as:
Bartels, M., Rietveld, M.J.H.,Van Baal, G.C.M., and Boomsma, D.I. (2002). Genetic and Environmental Influences on the Development of Intelligence. *Behavior Genetics* , 32(4), 237-249.

ABSTRACT

Measures of intelligence were collected in 209 twin pairs at 5, 7, 10, and 12 years of age, as part of a longitudinal project on intelligence, brain function and behavioral problems. Intelligence was measured at 5, 7, and 10 years of age with the RAKIT, a well-known Dutch intelligence test, consisting of 6 subscales. At 12 years of age the complete WISC-R was administered (12 subscales). Both intelligence tests resulted in a measure of Full-Scale IQ (FSIQ). Participation-rate is around 93% at age 12. Correlation coefficients over time are high: ($r_{(5-7)}$= .65; $r_{(5-10)}$= .65; $r_{(5-12)}$=.64; $r_{(7-10)}$=.72; $r_{(7-12)}$=.69 and $r_{(10-12)}$=.78). Genetic analyses show significant heritabilities at all ages, with the expected increase of genetic influences and decrease of shared environmental influences over the years. Genetic influences seem to be the main driving force behind continuity in general cognitive ability, represented by a common factor influencing FSIQ at all ages. Shared environmental influences are responsible for stability as well as change in the development of cognitive abilities, represented by a common factor influencing FSIQ at all ages and age-specific influences, respectively.

INTRODUCTION

Heritability of intelligence has been studied extensively, both in adults and in children, but far less is known about the developmental genetics of cognitive abilities. Many behavior genetic studies yield the largely consistent result that genetic differences account for at least 50% of the observed variability in cognition in adults (e.g. Bouchard and McGue, 1981; McCartney *et al*., 1990; Bratko, 1996; Rijsdijk *et al*., 1997, 1998; Alarcón *et al*, 1998, 1999, Posthuma *et al*., 2000). It is also well established that the genetic influences on cognitive functioning increase throughout development, whereas influences of common environment decrease (e.g. Skodak and Skeels, 1949; Wilson, 1983; Labuda *et al*., 1986; Fulker *et al*., 1988; Loehlin *et al*., 1989; McCartney *et al*., 1990; McGue *et al*., 1993; Boomsma, 1993; Plomin *et al*., 1997; Boomsma and Van Baal, 1998; Alarcón, 1998, 1999). A few longitudinal studies have focused on the influences of genes and environment on cognitive development rather than cognition at specific ages. New genetic influences at different ages and a common factor for shared environmental influences have been found (Colorado Adoption Project; e.g. Plomin and DeFries, 1985; Louisville Twin Study; e.g. Wilson, 1983; Eaves *et al*., 1986).

Longitudinal twin and family data allow the study of persistence and change of genetic, shared environmental and nonshared environmental influences. The genetic and environmental influences may exert their effects following several possible mechanisms. First, genetic or environmental factors may exert a continuous influence from their time of onset (common factor influences). This mechanism implies that the same genetic or environmental factors are responsible for stability, possibly with age-dependent factor loadings. Second, genetic and environmental influences may be specific at a certain age and exert an effect on cognition at that age only. Change in cognitive development may be due to these age specific factors. Finally, there can be a simplex-like continuity in genetic and environmental effects (Eaves *et al*., 1986; Boomsma and Molenaar, 1987). In this simplex-like continuity, there are effects specific to each age and there are 'carry-over effects' or transmission effects from one age to the subsequent age (Figure 7.1.). In other words, earlier influences may be transmitted from one occasion to the next and new influences (innovations) may come into play at each occasion. Data that are collected from the same subjects repeatedly in time often display this simplex structure for the observed correlations among the measures at different time points. Specifically, it is observed that correlations are highest among adjoining occasions and that they decrease systematically as the distance between time points increases (Guttman, 1954).

Notable longitudinal studies on cognition are the Colorado Adoption Project (CAP) (e.g. Plomin and DeFries, 1985) and the Louisville Twin Study (LTS) (e.g. Wilson, 1983). These studies are more or less comparable to the current study, in which intelligence is

assessed longitudinally in twins from age 5 to age 12. We will introduce CAP and LTS and also mention other studies, which offer an insight in the genetic and environmental patterns that account for variance in cognitive development (for reviews, see also Thompson, 1993; Patrick, 2000).

The CAP is a longitudinal 'full' adoption study of behavioral development. The study started in 1975 and included adopted children and their adoptive and biological parents. Children in the sample were tested yearly on age-appropriate cognitive measures. Until now, longitudinal results from 1 to 16 years of age have been published for the CAP. The CAP original sample consisted of 245 adoptive families and 245 nonadoptive control families. In 1999 the CAP sample consisted of 129 adopted individuals tested at 16 years of age and their adoptive and biological parents. The nonadoptive (control) sample included 125 sets of parents and nonadoptive children (Alarcón *et al*., 1999).

The LTS was initiated in 1957 by Falkner. In the LTS, twins were tested every three months in the first year of life. Testing continued at 6-months intervals during second and third year of life, and annually through age 9 with follow-up visits at age 15 and adulthood. In 1983, the sample of the LTS consisted of 494 pairs of twins active in the longitudinal study, ranging in age from 3 months to 15 years. Recruitment has been an ongoing process, with 25-35 pairs added each year since 1963. However, like in every longitudinal design, the study suffers from dropouts over the years.

Sophistication in developmental behavior genetics involves the formulation of models that attempt to describe the etiology of genetic and environmental influences on variation in cognitive development. Phillips and Fulker (1989) developed a model, based on a quasi-simplex model presented earlier by Eaves *et al.* (1986), in which it was possible to distinguish between the three possible longitudinal mechanisms (time-specific, common factor, simplex). This model was applied to a large data set combined from several major projects (Cardon *et al*., 1992). CAP data at ages 1, 2, 3, 4, and 7 were combined with data from twins at ages 1, 2, and 3 from MacArthur Longitudinal Twin Study (MLTS) (Plomin *et al*, 1990) and the Twin Infant Project (TIP) (DiLalla *et al*., 1990; Benson, *et al*., 1993). The best model for genetic influences on IQ was a simplex model, with time-specific innovations included. That is, genetic variation initially shown at age 1 is expressed through at least age 7 with new genetic variation, independent of the initial genetic influence, at ages 2,3, and 7, but not at age 4. For shared environment the best-fitting model showed only a single common factor influence on IQ, with equal factor loadings at each age. This longitudinal outcome suggests that shared environmental effects contribute to continuity only. In complete contrast is the picture that emerged for nonshared environment. For nonshared environment the influences were specific to each time-point

which implied that change in cognition is, at least partly, accounted for by these influences.

In a subsequent publication involving CAP, TIP, and MLTS subjects, including subjects from the CAP sample at age 9, very similar results were found (Fulker *et al.*, 1993a). This time, a Cholesky decomposition was used. A common genetic factor present at year one continued to account for observed variance in IQ, but with diminishing impact with increasing age. Evidence for genetic change at two important developmental transitions was found. The first was transition from infancy to early childhood (age 2 and 3). The second was the transition from early to middle childhood (age 7). Fulker and colleagues (1993a) speculated that the new genetic influence at age 7 might be in response to the 'novel environmental challenge' of schooling. No new genetic effect was apparent at age 9. Further, there was one continuous source of shared environmental influence across all ages. Application of the quasi-simplex model to the same data yielded identical results (Cherny and Cardon, 1994). Finally, the only longitudinal model-fitting results based on LTS data showed that a simplex model gave a better fit compared to a common factor model for genetic effects from ages 1 to 9 years (Humpreys and Davey, 1988).

To summarize, the general picture that emerges from these studies with young children is that genetic effects account both for stability and change in cognitive performance. This is implied by the simplex-like structure with time-specific innovation effects. Shared environmental effects appear to account for stability in intellectual performance, indicated by the single common factor structure, without age-specific effects. Consistent across studies, the nonshared environment is best modeled as exerting time specific influences only. This structure implies that the nonshared environment is important in explaining variance in cognitive performance at each age, but not in explaining stability of cognitive performance across ages.

In the present longitudinal study structural modeling techniques were used to examine the influences of genetic and environmental factors on development of Full-Scale IQ (FSIQ), using data of 209 Dutch twin pairs tested at 5, 7, 10, and 12 years of age. In addition to estimating the importance of heritability and environmental influences, the focus was on the developmental pattern of cognition. A genetic simplex model and a common factor model were used to study continuity and changes of genetic and environmental influences over time. Based on previous results of longitudinal studies on the development of cognitive functioning, a simplex structure for genetic influences was expected. Further, it was assumed that shared environmental factors show continuing effects over the years and nonshared environmental influences are age specific only.

METHODS

Participants

This study is part of an ongoing, longitudinal study of the development of intelligence and problem behavior. The study started in 1992 with recruitment of 209 twin pairs from the Netherlands Twin Register (NTR; Boomsma *et al.*, 1992; Boomsma, 1998; Boomsma *et al.*, 2002). The initial sample of 209 twin pairs was selected on the basis of age and zygosity of the twins, and their city of residence. Mean age at the first measurement occasion was 5.3 years (80% ranging from 5 years and 1 month to 5 years and 6 months). At the second measurement occasion mean age was 6.8 years (80% ranging from 6 years and 6 months to 7 years and 1 month). Mean age at the third measurement occasion was 10 years (80% ranging from 9 years and 11 months to 10 years and 1 month). Mean age at the fourth measurement occasion was 12 years and several day's (80% ranging from 11 years and 11 months to 12 years and 1 month). Zygosity of the same-sex twins was established by either blood group polymorphisms (137 pairs) or DNA analyses (24 pairs), and in a few pairs by physical resemblance assessed by the test-administrator (9 pairs). There were 47 monozygotic female (MZF), 37 dizygotic female (DZF), 42 monozygotic male (MZM), 44 dizygotic male (DZM), and 39 dizygotic pairs of opposite sex (DOS). The intelligence test was administered to all 209 twin pairs at age 5. At the second measurement occasion (age 7) 192 pairs of the original sample provided complete data on all subtests. The number of participating twin pairs increased to 197 when the children were tested around their 10th birthday. At the fourth measurement occasion (age 12) 192 twin pairs participated. A small group of four families refused consistently to participate after the first measurement occasion. Five families dropped out at both ages 10 and 12. The remaining nonparticipants refused participation at one measurement occasion. At age 5 and 12, one incomplete twin pair can be found in the data because of difficulties during testing (age 5) and refusal to participate (age 12). Due to serious loss of hearing one twin pair was assigned missing value at all four ages for FSIQ. This left a sample of 176 twin pairs with complete data at all four ages. No significant difference in initial FSIQ (at age 5) has been found for twins who dropped out on one or more of the following occasions ($F_{3, 415} = 2.25$, $p = .082$). Details on the demographic characteristics of the sample and information on parental occupation can be found in Rietveld *et al.*, (2000).

Procedure and intelligence tests

At ages 5 and 7 years the twins participated in a study on the development of cognitive abilities and brain-activity (Boomsma and Van Baal, 1998). At both measurement occasions the twin and their family visited the laboratory at the university. While one of the twins participated in the electro-physiological experiment, the co-twin participated in

an intelligence test. At age 10 and 12 years a different procedure was followed. The twins and their parents could choose whether they preferred to come to the university or whether they preferred to be visited at home to participate in the intelligence test. The majority of the families (around 70% at both ages) preferred testing at home. No significant difference in FSIQ was observed between children tested at home or at the university. The intelligence test was assessed by an experienced test-administrator. At ages 5, 7, and 10 the test took approximately one hour to complete and at age 12 the test took one and a half-hour to complete. All children received a present afterwards.

At age 5,7, and 10 the children were tested with the Revised Amsterdamse Kinder Intelligentie Test (RAKIT) (Bleichrodt *et al*, 1984). Six subtests, with age-appropriate items, were employed to assess cognitive functioning. Raw subtest total scores are corrected for age and transformed into standardized scores with a mean of 15 and a standard deviation of 5. The total IQ score is based on the combination of these transformed subtests with a mean of 100 and standard deviation of 15. The standardization is based on a population sample of Dutch 6 to 11-year-old children. No difference is made for boys and girls. For further details on this well-known Dutch intelligence test see Rietveld *et al.* (2000). At age 12 the twins conducted the complete version of the WISC-R, Dutch version (Van Haasen *et al*., 1986). The WISC-R consists of 12 subtests, 6 mainly verbal and 6 mainly non-verbal. The subtest scores are standardized, based on results of same-aged children in the Netherlands. No differences are made for boys and girls. Addition of the twelve standardized subtest scores results in FSIQ. The concurrent validity of the RAKIT and the WISC-R is .86 (Pijl *et al*., 1984).

Statistical analyses

Descriptive statistics for FSIQ were calculated using SPSS/windows 10. Twin correlations with their 95% confidence intervals at each age have been calculated. These correlations are informative on the importance of genes and environment in explaining observed variance at each age. To assess stability of intelligence, phenotypic cross correlations over time were calculated. MZ and DZ cross correlations over time have been calculated to get a first impression of the genetic and environmental contributions to the covariance over time.

Genetic Modeling

Univariate model fitting procedures were used to estimate genetic and environmental influences at each age separately and to investigate the presence of sex-differences and influences of sex-specific genes in these data. Genetic model fitting of twin data allows for separation of the observed phenotypic variance into its genetic and environmental

components. Additive genetic variance (A), is the variance that results from the additive effects of alleles at each contributing genetic locus. Shared environmental variance (C) is the variance that results from environmental events common to both members of a twin pair. Nonshared environmental variance (E) is the variance that results from environmental effects that are not shared by members of a twin pair. Estimates of the nonshared environmental effects also include measurement error. To account for this source of variance, E is always specified in the model.

The different degree of genetic relatedness between monozygotic (MZ) and dizygotic (DZ) twin pairs was used to estimate the contribution of these factors to the phenotypic variation in cognitive abilities. Similarities for MZ twins are assumed to be due to additive genetic influences plus environmental influences that are shared by both members of a twin pair. Experiences that make MZ twins different from one another are nonshared environmental influences. Because DZ twins share 50% of their genetic material on average, like other siblings, genetic factors contribute only half to their resemblance. As for MZ twins the shared environment contributes fully. Model fitting to twin data is based on the comparison of the variance-covariance matrices in MZ and DZ twins. Exploiting the known difference in genetic contribution to intra-pair resemblance of MZ and DZ twin pairs, influences of additive genetic, shared environmental and nonshared environmental factors are estimated using the computer program Mx (Neale *et al.*, 1999).

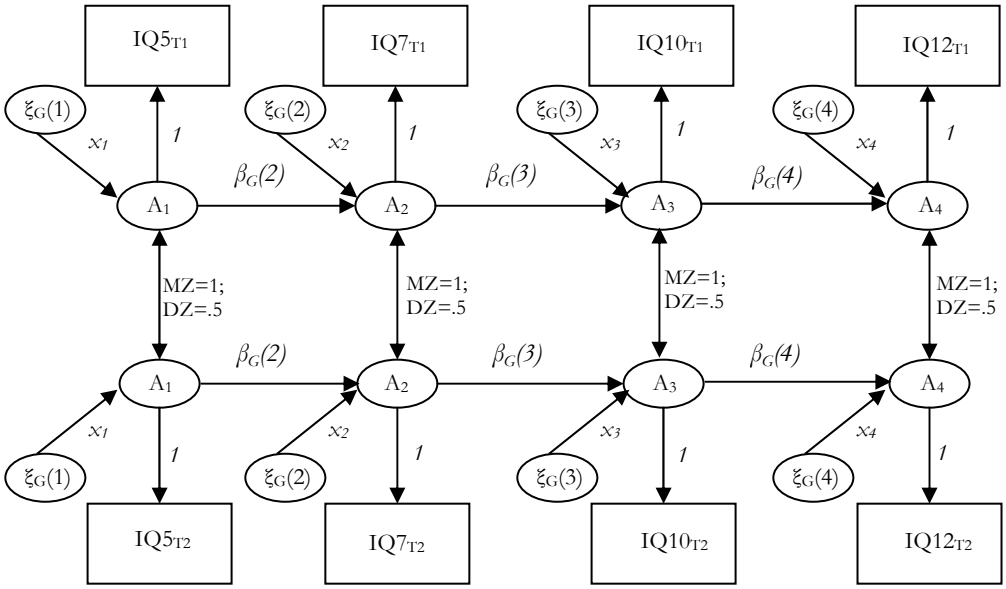
Differences between boys and girls can occur in two ways. First a difference in the magnitude of additive genetic, shared environmental and nonshared environmental influences can exist, represented in a distinct pattern of twin correlations for boys and girls. Second, heterogeneity, an expression of different genes in boys and girls, can occur. This heterogeneity would be represented by a lower twin correlation in dizygotic twins of opposite sex in comparison to dizygotic same sex twins. Differences in magnitude of additive genetic, shared environmental and nonshared environmental influences is tested by the change in fit after constraining the parameter estimates equal for boys and girls. Testing for heterogeneity is accomplished by testing the genetic correlations between two members of a dizygotic twin of opposite sex. Normally the genetic correlation of DZ twins is fixed at .5. Heterogeneity would result in a genetic correlation of less than .5.

Multivariate genetic model fitting techniques were used to obtain insight in the developmental pattern of cognitive functioning and to obtain estimates of the genetic and environmental influences on cognitive development. Parameters were estimated by maximum likelihood, using the computer program Mx (Neale *et al.*, 1999). Rather than decomposing the variance of a measurement into genetic and environmental sources of variance, multivariate genetic analysis decomposes the variance of each measurement occasion and the covariance between the measurement occasions into genetic and

environmental sources. The total variances and covariances were decomposed into additive genetic (A), shared environmental (C), and nonshared environmental (E) parts. First, to get an initial insight in the variance and covariance structure a Cholesky decomposition model was applied to the data. Next, to investigate the stability and change in FSIQ a simplex model was applied to the data. For each source of variance (A,C, and E) a simplex structure was specified.

Figure 7.1.

A Simplex Model in which IQ5….IQ12 represent the observed IQ at measurement occasion one to measurement occasion four for twin 1 (T1) and twin 2 (T2). $\xi_G(1)$… $\xi_G(4)$ represent the innovations and $\beta_G(2)$… $\beta_G(2)$ the transmission effects.



A simplex model is a first order auto-regressive process. In the simplex model covariances among the four ages of measurement are specified by genetic and environmental factors specific to each age and by 'carry-over effects' or transmission of these factors to subsequent ages. The model specifies the variance nonshared to each measurement occasion by an innovation term that comes into play at each time point. The variance is a product of the age specific effects and age-to-age transmission effect (see Appendix 7.1 and Figure 7.1). Finally, it was investigated whether a common factor,

possibly with age-dependent factor loadings and age specifics influences, could replace the simplex structure for genetic and shared environmental influences.

To make optimal use of all available data, analyses were performed on the raw data. Submodels were compared by hierarchic χ^2 tests. The χ^2 statistic is computed by subtracting –2LL for the full model from that for a reduced model ($\chi^2 = 2\ (LL_1 - LL_0)$). A good model is indicated by a low non-significant χ^2 test statistic ($P > .05$). Apart from the χ^2 test statistic, Akaike's Information Criterion (AIC = χ^2 - 2 × degrees of freedom) was computed. The lower the AIC the better the fit of the model to the observed data.

Reductions of the model were based on the expectations raised by previous studies. In detail, a simplex structure for genetic influences, a common factor for shared environmental influences and time-specific structure for nonshared environmental influences is expected. Estimates of genetic, shared environmental and nonshared environmental influences on the age specific variance and between age covariance of general cognitive abilities are reported based on the Cholesky decomposition model, the full simplex model and the best fitting reduced model.

Table 7.1.

Descriptive statistics for Full-Scale IQ at different ages

	N[a]	Mean Age	Min	Max	Mean	Std	Skewness		Kurtosis	
								s.e.		s.e.
FSIQ5	415	5.3	64	142	102.75	13.18	-.059	.120	.209	.239
FSIQ7	382	6.8	62	145	102.90	14.67	-.127	.125	.023	.249
FSIQ10	392	10.0	63	145	106.96	15.54	-.066	.123	-.166	.246
FSIQ12	381	12.0	61	138	100.03	13.18	-.039	.125	.177	.249

[a] number of children in the study

RESULTS

Descriptive statistics for FSIQ at 5, 7, 10, and 12 years of age showed that the variables were approximately normal distributed (Table 7.1). Table 7.2 shows the twin correlations for the five zygosity groups calculated separately for each age. MZ correlations are higher than DZ correlations suggesting genetic influences at each age. The low DOS correlation at age 12 suggests heterogeneity and univariate model fitting procedures were used to explore this possibility. Estimates for genetic and shared environmental influences based on the univariate model-fitting procedure are presented in Table 7.3. These results are consistent with previous results (Boomsma and Van Baal, 1998; Bouchard and McGue, 1981) showing increase of genetic influences and diminishing effects of shared

environment over the years. Shared environmental influences are insignificant at ages 10 and 12. Univariate model fitting showed no presence of sex-differences at the four ages separately and no presence of sex specific genes at age 12.

To get a first impression of the developmental pattern of cognitive abilities, phenotypic cross-correlations over time were calculated (Table 7.4). All correlations are rather large, which indicates a strong degree of stability of intellectual performance. This structure may best be described by a common factor mechanism.

Cross correlations over time for monozygotic (MZ) and dizygotic (DZ) twins were calculated separately to explore the genetic and environmental influences on the observed stability. As can be seen in Table 7.4, the MZ cross correlations over time (above the diagonal) are higher than the DZ cross correlations over time (below the diagonal), suggesting that stability in intelligence over time is mainly due to genetic factors. Further, when the correlations of the adjoining age-intervals are compared (age 5 to 7; age 7 to 10; age 10 to 12) the increased difference between MZ and DZ correlations suggests an increase in the genetic contribution to stability with increasing age.

Table 7.2.

Twin Correlations for FSIQ with 95% confidence intervals

	MZF[a]	**DZF**	**MZM**	**DZM**	**DOS**
FSIQ5	.78 *(.64-.87)* *46[b]*	.73 *(.53-.85)* *37*	.77 *(.62-.87)* *42*	.53 *(.29-.72)* *43*	.64 *(.41-.79)* *39*
FSIQ7	.77 *(.61-.87)* *41*	.50 *(.20-.70)* *34*	.56 *(.29-.74)* *37*	.41 *(.13-.63)* *41*	.56 *(.30-.74)* *38*
FSIQ10	.87 *(.78-.92)* *43*	.45 *(.16-.67)* *37*	.73 *(.54-.85)* *38*	.53 *(.28-.72)* *41*	.50 *(.21-.70))* *37*
FSIQ12	.86 *(.76-.92)* *43*	.67 *(.46-.82)* *37*	.84 *(.71-.91)* *36*	.57 *(.32-.75)* *39*	.35 *(.03-.60)* *35*

[a] MZF= monozygotic female, DZF= dizygotic female, MZM= monozygotic male, DZM=dizyotic males, DOS= dizygotic opposite sex; [b] number of complete twin pairs

Analyses were continued with the application of the different models to the longitudinal data. Model-fitting procedures yielded the results presented in Table 7.5. The simplex model without restrictions (model 2) was taken as a reference for evaluating changes in χ^2 and associated degrees of freedom of more parsimonious models. First, reduction of the model was based on the expectation of age-specific nonshared environmental factors only (model 3). No significant change in χ^2 arose.

Second, model reduction was based on the expectation of a common factor for shared environmental influences (model 4). Because the order of model reduction may influence the fit of the model, a model with a common factor for genetic influences and a simplex structure for shared environmental influences was fitted to the data as well (model 5). No clear distinction could be made between models 4 and 5, both being more parsimonious than model 3 but not significantly different.

Table 7.3.

Univariate Model-Fitting results and parameter estimates for FSIQ at the four ages.

	MODEL	-2LL	df	χ^2	df	p	A	C	E
FSIQ5	ACE + sd[a]	3178.20	408						
	ACE	3180.40	411	2.20	3	.53	.26 *(.03-.52)*	.50 *(.26-.68)*	.24 *(.18-.33)*
	AE	3193.71	412	13.31	1	.00			
	CE	3185.25	412	4.85	1	.03			
FSIQ7	ACE + sd	3051.33	375						
	ACE	3054.37	378	3.04	3	.39	.39 *(.07-.72)*	.30 *(.00-.55)*	.31 *(.23-.44)*
	AE	3058.12	379	3.75	1	.05	.70 *(.60-.78)*	-	.30 *(.22-.40)*
	CE	3059.83	379	5.46	1	.02			
FSIQ10	ACE + sd	3135.48	385						
	ACE	3140.87	388	5.39	3	.15	.54 *(.28-.83)*	.25 *(.00-.48)*	.21 *(.15-.29)*
	AE	3143.62	389	2.75	1	.10	.80 *(.72-.85)*	-	.20 *(.15-.28)*
	CE	3156.99	389	16.12	1	.00			
FSIQ12	ACE+ rg_{free}[b]	2903.40	373						
	ACE+ sd	2903.71	374	.31	1	.58			
	ACE	2908.72	377	5.01	3	.17	.64 *(.40-.88)*	.21 *(.00-.43)*	.15 *(.11-.22)*
	AE	2910.81	378	2.09	1	.15	.85 *(.79-.89)*	-	.15 *(.11-.21)*
	CE	2936.28	378	27.56	1	.00			

[a] model with sex differences for parameter estimates; [b] model with sex specific genes

A model with a common factor for both genetic and shared environmental influences, allowing for time specific influences as well, did not gave a significant worse fit (model 6). Further, it was tested whether dropping the age-specific influences, either genetic or shared environmental, altered the χ^2 significantly (model 7 and model 8). Based on the difference in χ^2 and the lower AIC, model 7 was preferred above model 8. The genetic (co)variance is modeled as a common factor without specifics, whereas the shared environmental (co)variance is modeled as a common factor with specifics. These results

suggest that stability in cognitive performance is mainly due to genetic factors. Finally, a model with a common factor, without time specific influences for both genetic and shared environmental influences showed a significant increase in χ^2 (model 9).

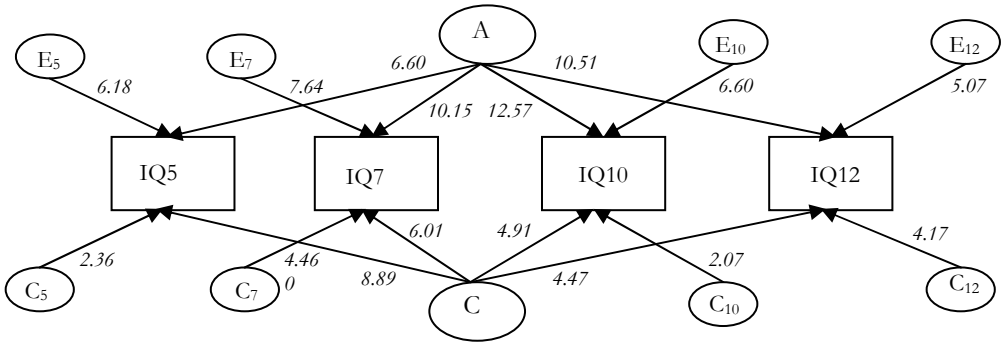


Figure 7.2.

Model 7: Common factor for additive genetic influences, common factor with age specific influences for shared environmental influences, and age specific influences for nonshared environmental factors.

Estimates of the path-coefficients for the best fitting model (model 7) are presented in Figure 7.2. Percentage of age specific variance explained by genetic, shared environmental and nonshared environmental factors based on the Cholesky decomposition (model 1), the full simplex model (model 2), and the best fitting model (model 7) are presented in table 7.6. Table 7.7 contains the percentages of between age covariances explained by genetic, shared environmental and nonshared environmental factors based on the Cholesky decomposition, the full simplex model, and the best fitting model. Indicated by the observed MZ and DZ cross-correlations, genes become more important in explaining stability in cognitive performance with increasing age. As opposed to this outcome, the shared environment accounts for a decreasing portion of the covariance between age-intervals, whereas the nonshared environment explains, on general, around one quarter of the total variance at each age (Table 7.6). The nonshared environment contributes only minimal to the observed covariance between ages (Table 7.7). It should be noted that the genetic, shared environmental and nonshared environmental variance components estimated from fitting the multivariate models to these data are somewhat different from what univariate analyses at each age separately might yield. These differences arise because the multivariate models take into account the cross-sibling cross-time covariance structure, which can impact the within-time parameter estimates. In addition, multivariate model-

fitting increases the power to detect shared environmental influences as a source of familial aggregation.

Table 7.4.

Phenotypic cross correlations for FSIQ, calculated for the complete dataset and MZ (above diagonal) and DZ (below diagonal) cross correlations over time for FSIQ.

Total Sample	**5**	**7**	**10**	**12**
5	1.00	.65 *(.59-.70)*	.65 *(.59-.70)*	.64 *(.57-.69)*
7		1.00	.72 *(.67-.77)*	.69 *(.63-.74)*
10			1.00	.78 *(.74-.82)*
12				1.00
DZ/MZ	**5b**	**7b**	**10b**	**12b**
5a	-	.66 *(.54-.75)*	.67 *(.56-.76)*	.68 *(.57-.77)*
7a	.42 *(.30-.54)*	-	.71 *(.60-.79)*	.68 *(.57-.77)*
10a	.42 *(.30-.54)*	.39 *(.26-.52)*	-	.79 *(.70-.85)*
12a	.42 *(.29-.54)*	.42 *(.29-.54)*	.45 *(.32-.57)*	-

DISCUSSION

The influences of genes and environment on cognitive development and on its developmental structure were studied in a longitudinal sample of Dutch twins at 5, 7, 10, and 12 years of age. It can be concluded that the development of general cognitive abilities is a continuous process. Continuity is represented by a common factor, with age specific factor loadings, for both genetic and shared environmental influences. Change in development, represented by age specific factors, are presented in the shared environmental structure and, as expected, in the nonshared environmental structure. Further, decomposition of the between age covariances in additive genetic, shared environmental, and nonshared environmental influences showed that the continuity in cognitive abilities is mainly due to additive genetic factors.

In this study increasing additive genetic influences and decreasing influences of shared environmental factors are found in both age specific variances and between age covariances. The increase of genetic influences on cognitive functioning throughout development is already well established in US samples and is now also found in a sample of Dutch twins.

Table 7.5. *Model-fitting results for FSIQ*

MODEL	-2LL	df	χ^2	df	c.t.m.	p	AIC
1. **A**: Cholesky **C**: Cholesky **E:** Cholesky	11527.470	1520					
2. **A**:full simplex structure **C**:full simplex structure **E**:full simplex structure with time-specific factors[a]	11510.105	1508					
3. **A**:full simplex structure **C**:full simplex structure **E**:time-specific factors only[b]	11517.278	1512	5.681	4	2	.22	-2.319
4. **A**:full simplex structure **C**:common factor + specifics **E**:time-specific factors only[b]	11513.722	1511	3.617	3	2	.31	-2.383
5. **A**:common factor + specifics **C**:full simplex structure **E**:time-specific factors only[b]	11517.303	1511	7.198	3	2	.07	1.198
6. **A**:common factor+specifcs **C**:common factor + specifics **E**:time-specific factors only[b]	11513.743	1510	3.638	2	2	.16	-.362
7. **A:common factor** **C:common factor + specifics** **E:time-specific factors only[b]**	**11513.743**	**1514**	**-**	**4**	**6**	**1.00**	**-8.000**
8. **A**:common factor +specifics **C**:common factor **E**:time-specific factors only[b]	11521.677	1514	7.934	4	6	.09	-.066
9. **A**:common factor **C**:common factor **E**:time-specific factors only[b]	11545.797	1518	32.05	8	6	.00	16.050

[a] these time specific factor are equal at all ages; [b] these time specific factors are estimated separately at every age

In the common factor pattern for genetic influences (model 7; Figure 7.2), increasing influences of heritability are represented by increasing factor loadings from age 5 to 10.

Further, in the common factor pattern for shared environmental influences, decreasing influences are represented by decreasing factor loadings from age 5 to 10 and decreasing age specific influences from age 5 to 10.

Table 7.6.

Percentage of variance explained by A,C en E, based on a Cholesky decomposition, a simplex model and the best fitting model, with 95% confidence intervals

Model 1	**Var**	**A cholesky**	**C cholesky**			**E cholesky**
	5	.30 *(.15-.54)*	.46 *(.24-.62)*			.24 *(.18-.30)*
	7	.42 *(.21-.67)*	.28 *(.05-.49)*			.30 *(.22-.39)*
	10	.61 *(.36-.82)*	.19 *(.00-.43)*			.20 *(.15-.27)*
	12	.62 *(.39-.86)*	.23 *(.00-.45)*			.15 *(.11-.21)*
Model 2	**Var**	**A simplex**	**C simplex**			**E simplex**
	5	.38 *(.22-.60)*	.39 *(.18-.56)*			.23 *(.18-.30)*
	7	.38 *(.20-.63)*	.32 *(.09-.50)*			.30 *(.23-.39)*
	10	.72 *(.33-.83)*	.08 *(.00-.46)*			.20 *(.15-.27)*
	12	.62 *(.37-.85)*	.23 *(.01-.46)*			.15 *(.11-.21)*
Model 7	**Var**	**A common**	**C total**	**common**	**specific**	**E time specific**
	5	.26 *(.14-.46)*	.51 *(.31-.65)*	.47	.04	.23 *(.18-.29)*
	7	.47 *(.29-.62)*	.26 *(.11-.45)*	.17	.09	.27 *(.21-.35)*
	10	.69 *(.49-.82)*	.12 *(.00-.33)*	.10	.02	.19 *(.14-.25)*
	12	.64 *(.45-.77)*	.21 *(.09-.40)*	.11	.10	.15 *(.11-.20)*

The developmental pattern for genetic influences found in this study is partly different from previous, comparable studies like the combined study of CAP, MLTS, and TIP (Cardon *et al.*, 1992; Fulker, *et al.*, 1993). Results provided by these studies show a simplex pattern for genetic influences with genetic innovation at 2, 3, and 7 years of age, with the suggestion that genetic innovation at age 7 may be due to 'the novel environmental challenge of schooling'. In our study no indication for genetic innovation is obtained.

Comparison of the different longitudinal studies is limited due to distinct ages of testing. In our study no information is available for cognitive development prior to age 5 and the results, mainly presented by the CAP studies, provide no information on the development of general cognitive ability between age 9 and 16. Another difficulty in longitudinal studies in general and in comparing different longitudinal studies on cognitive development in particular is the measurement of cognitive performance. There are no cognitive assessments that are common to all ages, so different age appropriate

instruments must be used. One of the difficulties that comes along with this issue is that no distinction can be made between true changes in development and changes due to different measurement instruments.

Table 7.7.

Percentage of covariance explained by A, C en E, based on a Cholesky model, a simplex model and a restircited models with their 95% confidence intervals

Model 1	**Covariance**	**A cholesky**	**C cholesky**	**E cholesky**
	5-7	.55 *(.32-.87)*	.40 *(.10-.63)*	.05 *(.00-.12)*
	5-10	.66 *(.40-.99)*	.34 *(.01-.58)*	.00 *(.00-.07)*
	5-12	.66 *(.42-.98)*	.34 *(.02-.58)*	.00 *(.00-.04)*
	7-10	.71 *(.43-.93)*	.21 *(.00-.48)*	.08 *(.02-.15)*
	7-12	.73 *(.46-.98)*	.24 *(.01-.50)*	.03 *(.00-.10)*
	10-12	.78 *(.51-.99)*	.20 *(.00-.45)*	.02 *(.00-.08)*
Model 2	**Covariance**	**A simplex**	**C simplex**	**E simplex**
	5-7	.59 *(.33-.95)*	.38 *(.04-.63)*	.03 *(.00-.10)*
	5-10	.82 *(.55-.99)*	.17 *(.00-.45)*	.01 *(.00-.04)*
	5-12	.76 *(.48-.99)*	.24 *(.01-.52)*	.00 *(.00-.02)*
	7-10	.73 *(.46-.97)*	.20 *(.00-.46)*	.07 *(.01-.15)*
	7-12	.69 *(.42-.98)*	.29 *(.01-.56)*	.02 *(.00-.07)*
	10-12	.84 *(.55-.99)*	.13 *(.00-.42)*	.03 *(.00-.08)*
Model 7	**Covariance**	**A common factor**	**C common factor**	**E time specific[a]**
	5-7	.56 *(.35-.83)*	.44 *(.17-65)*	-
	5-10	.66 *(.42-.95)*	.34 *(.05-.58)*	-
	5-12	.64 *(.41-.92)*	.36 *(.08-.59)*	-
	7-10	.81 *(.55-.98)*	.19 *(.02-.45)*	-
	7-12	.80 *(.54-.98)*	.20 *(02-.46)*	-
	10-12	.86 *(.63-.99)*	.14 *(.01-.37)*	-

[a] E is represented in time specific influences only.

A major advantage of our longitudinal study is that the same intelligence test (RAKIT), with age specific items, is used at the first three measurement occasions. Further, the intelligence test used at the fourth measurement occasion (WISC-R) shows a high concurrent validity with the RAKIT (.86 for full-scale IQ) (Pijl *et al.*, 1984).

More striking is the finding for shared environmental influences. Previous studies suggested a common factor for shared environmental influences. Our study indicates that,

besides a continuing influence of shared environmental factors, age specific influences are present. These age-specific effects were significant but the proportion of variance explained is much smaller compared to the proportion explained by the shared environmental factor common to all ages. This common factor could be accounted for by SES and parental education, as these environmental aspects are not sensitive to large changes over a time-span of 7 years.

Aspects outside the family environment, like friends or being a member of a sportsclub, might also cause similarities between two children of a twin pair during childhood. For the age specific shared environmental influences one may consider the school environment. Information on same or different teacher for both children of a large sample of 12 year old twin pairs (n=1164) indicates that in 63% of the cases both children of a twin pair are taught by the same teacher, whereas 37% go to separate classes. This ratio makes teacher or classroom environment a shared environmental influence for the majority of the children. Since, in the Dutch school system children move to a different teacher each school year, this results in a lack of continuity in this particular aspect of shared environment. So, these shared but age-specific experiences within the class-room may be represented by the age-specific factors as specified significant in the best fitting model. In addition to SES and aspects of school, the direct neighborhood, experienced during childhood, may contain shared environmental influences. Nearly half of the initial sample (47%) changed residency after the twins were born. The majority of those families who moved once during the twins' lifetime did so before the twins' fifth birthday. If this particular source of shared environmental variance has an impact on the development of cognition, it may be considered a continuous source of influence. That is, the impact of change of residency remains detected years after the actual change took place. Further, the model fitting results imply that this hypothetical influence of the shared environment diminishes with increasing age. This is what one expects when a change of domicile has taken place early in a child's life. The nonshared environment was found to explain a substantial portion of the variance at each age (best model, range from 15% to 27%). With respect to developmental aspects of the data the nonshared environment acts in a well-established manner. The environment that is uniquely experienced by an individual contributes to change rather than stability in cognitive performance.

The above mentioned findings are in line with those obtained by multivariate analyses of the RAKIT subtests collected at the twin's age 5, 7 an 10 (Rietveld *et al.*, submitted). Subtest performance, either verbal or nonverbal, displays stability mainly due to genetic effects and to a lesser extent to shared environmental effects. The nonshared environment is important at each age, but plays no role of significance when one attempts to explain stability in subtest performance.

For future purposes of disentangling genetic and environmental influences on cognitive development it might be important to collect more information on possible shared or nonshared environmental influences as previous studies on the development of cognition mainly focus on heritability estimates. The nature of the influences of shared and nonshared environment is underexposed and only modestly discussed in behavior genetic literature. Ideally, one should measure a range of potential environmental influences to be able to gain more insight into the exact nature of these influences.

Besides the above mentioned use of different age appropriate tests, longitudinal studies in general are subjected to other unavoidable difficulties. A major difficulty in longitudinal studies is the participation rate. By studying the same subjects over the years dropout is inevitable. In our study the dropout rate is low. Over 90% of the initial sample continued to participate at the fourth measurement occasion and the reasons for leaving the study were found unrelated to the initial measurement of the twins FSIQ. Complete intelligence data at all ages are available for 84% of the sample.

Our ongoing longitudinal study has the potential to overcome some of the mentioned shortcomings of the unknown influence of the use of different intelligence tests. In order to clearify out the influences on the change of FSIQ test between age 10 and 12 a sample of younger siblings of the twins will be tested by making use of both tests. And, a continuity of the study has just started to see whether hormonal influences, induced by puberty, change the developmental pattern of general and specific cognitive abilities between age 12 and 14.

In summary, the results of our study did not fully reach our prior expectations based on previous studies. In our study genetic influences are the main driving force behind continuity in general cognitive ability. The shared environment contributes to continuity and to a lesser extent to change. As expected, nonshared environmental influences contribute to change of cognitive abilities solely.

Appendix 7.1.

The longitudinal simplex model exists of a measurement model, which represents the relation between the latent and the observed variables (FSIQ = F +U), in which FSIQ is the observed variable, F is the latent variable (A, C or E) and U is the measurement error. Further it exists of a structural equation model, which represents the relation among the latent variables ($A = \beta A + \xi$), in which A is the latent variable, β is the transmission factor and ξ is the innovation factor. A simplex model is a first-order auto-regressive process. In other words; each latent variable is influenced by the preceding latent variable (see Figure 7.1):

$$A_i = \beta_i \ A_{i-1} + \xi_i$$

In which β_i is the autoregressive (transmission) coefficient and ξ_i represent innovation at that point in time. Further, units of measurement in the latent variables are the same as in the observed variables resulting in:

$$FSIQ=(I-\beta)^{-1} * \xi + U$$

Hence, the expected additive genetic covariance matrix is:

$$H = (I-G)^{-1} * A*A' * ((I-G)^{-1})'$$

where genetic transmission parameters are modeled in matrix G, a 4*4 matrix with three transmission parameters on its subdiagonal, based on the four point in time used in this study. These autoregressive coefficients ($\beta_G(t)$) are a measure of the amount of genetic variation at time point *t*-1 that is transmitted to time point *t* and therefore associated with stability. Genetic innovation parameters are modeled in matrix A, a 4*4 diagonal matrix. These innovation parameters ($\xi_G(t)$) denote the effects of new genes turned on at time point *t* and will therefore lower the stability of the genetic process between *t*-1 and *t*. Similar parameter matrices can be defined for nonshared and common environment.

8

Heritability of Educational Achievement in 12-year-olds and the overlap with Cognitive Ability

This chapter is published as:
Bartels, M., Rietveld, M.J.H.,Van Baal, G.C.M., and Boomsma, D.I. (2002). Heritability of Educational Achievement in 12-year-olds and the overlap with Cognitve Abilites. *Twin Research*, 5(6), 1-10.

ABSTRACT

The aim of this study is to explore to what extent psychometric IQ and scholastic achievement, as assessed by the CITO high-school entrance test, are correlated. In addition, it was investigated whether this expected correlation is due to a common genetic background, shared or nonshared environmental influences common to CITO and intelligence or a combination of these influences. To this end multivariate behavior genetic analyses with CITO and IQ at ages 5, 7, 10 and 12 years have been conducted. The correlations were .41, .50, .60, and .63 between CITO and IQ assessed at age 5, 7, 10, and 12 respectively. The results of the analyses point to genetic effects as the main source of variance in CITO and an important source of covariance between CITO and IQ. Additive genetic effects account for 60% of the individual differences found in CITO scores in a large sample of Dutch 12-year-olds. This high heritability indicates that the CITO might be a valuable instrument to assess individual differences in cognitive abilities in children but might not be the right instrument to put the effect of education to the test.

INTRODUCTION

In the Netherlands, nowadays, much value is attached to the results of a national test of educational achievement (CITO), administered around age 12, in order to determine high school entrance level. The results of the test are often used as an independent judgement, besides the teachers' opinion, in advising the parents on future educational level of their child. From a historical perspective, this attention for 'independent' testing has to do with the possibilities for selection. The establisher of the CITO (Eindtoets Basisonderwijs 2002, Arnhem: Citogroep, 2002) emphasizes that this national test of educational achievement puts the effect of education in a particular school to the test besides measuring possible learning potential or cognitive abilities in children (Geldermans, 2001). It is hypothesized that success in scholastic achievement depends on the quality of the elementary school. A large number of articles in Dutch daily newspapers are dedicated to the influences of the school population and school neighborhood on the test results of the pupils. In these articles the influences of SES and ethnic background of the majority of the children at a certain school are considered important factors to classify the school and the future success of the pupils. Several studies agree on the claim that family variables (e.g. family size, SES, parental involvement, cultural level) influence the development and educational achievement of children (Christenson, *et al.*, 1992; Marjoribanks, 1994, Garcia and Rosel, 2001). If this is true, influences of shared environmental factors on CITO would show up as significant in the classical twin design. Alternatively, parental SES may reflect the parents' cognitive abilities. Heritable influences on cognition would predict CITO scores to be genetically mediated.

Intelligence has been found to explain a significant amount of the variance in educational achievement (Eaves and Darch, 1990, Jensen, 1972). In the Netherlands no attempt has been made to address the question about the nature of a possible association between the results of the CITO and cognitive abilities, as measured by psychometric IQ. Even more striking, no attention has been paid to possible genetic influences on the results of the national test of educational achievement. In emphasizing that the CITO is a test for the level of the school and the classification of children, the CITO-group may underestimate the true content and value of the results of this test. It could be interesting to sort out whether the possible association between intelligence and results of the CITO is based either on overlapping genetic influences, overlapping environmental influences (SES, school population), or both.

Numerous behavior genetic studies have been conducted in which cognition and educational achievement are examined separately. Studies on cognition yield the largely consistent result that genetic differences account for at least 50% of the observed variability in cognition in adults (e.g. Bouchard and McGue, 1981; McCartney *et al.*, 1990;

Bratko, 1996; Rijsdijk *et al.*, 1997, 1998; Alarcón *et al*, 1998, 1999; Posthuma *et al.*, 2000). It is also well established that the genetic influences on cognitive functioning increase throughout development, whereas influences of common environment decrease (e.g. Skodak and Skeels, 1949; Wilson, 1983; Labuda *et al.*, 1986; Fulker *et al.*, 1988; Loehlin *et al.*, 1989; McCartney *et al.*, 1990; McGue *et al.*, 1993; Boomsma, 1993; Plomin *et al.*, 1997; Boomsma and Van Baal, 1998; Alarcón, 1998, 1999, Bartels *et al.*, 2002). A few longitudinal studies have focused on the influences of genes and environment on cognitive development rather than cognition at specific ages. Results of these studies prove intelligence to be one of the most stable phenotypes (Colorado Adoption Project; e.g. Plomin and DeFries, 1985; Louisville Twin Study; e.g. Wilson, 1983; Eaves *et al.*, 1986; Netherlands Twin Register, Boomsma *et al.*, 1992; 2002; Boomsma, 1998). In the Dutch longitudinal sample this stability in intelligence seems to be mainly genetically mediated. Environmental factors give rise to stability as well as change of cognitive functioning over the years (Bartels *et al.*, 2002).

There are fewer behavioral genetic studies of scholastic achievement during childhood. However, genetic influences also seem to be an important factor explaining individual differences in achievement, but the influences of shared environmental factors can not be ruled out (Nichols, 1965; Martin and Martin, 1975; Martin, 1975; Loehlin and Nichols, 1976; Willerman *et al.*, 1977; Labuda *et al.*, 1986; for a review see Plomin, 1986; Thompson *et al.*, 1991).

Multivariate behavioral genetic models indicate that genetic effects are the primary source of variance underlying the phenotypic correlation between cognition and scholastic achievement (Petrill *et al.*, 1993; Wadsworth *et al.*, 1995a, 1995b). Thompson and colleagues (1991) showed genetic correlations between cognition and achievement tests ranging from .57 to .85., whereas shared environment correlations were essentially zero, and specific environment correlations were low (.00 to .19). Results of the Colorado Adoption Project, using related and unrelated sibling pairs, showed that genetic influences accounted for most of the phenotypic covariance among measures of cognitive ability (verbal comprehension and perceptual organization) and achievement (reading recognition and mathematics achievement), with much of the genetic covariation being due to influences shared with verbal ability (Wadsworth *et al.*, 1995a). An extension of the previous study by simultaneously analyzing parent-offspring and sibling data form the CAP yielded the same results (Wadsworth *et al.*, 1995b).

We have studied the development of cognitive abilities and the correlation with educational achievement in a large longitudinal sample of Dutch twins. A previous analysis on the heritability of cognition in this longitudinal sample of Dutch twins at ages 5, 7, 10 and 12, showed an increase in heritability over the years, ranging from 26% at age 5 to

85% at age 12. A decrease in shared environmental influences is observed. Shared environmental influences seem to be significant at age 5 and 7, but not at ages 10 and 12 (see table 8.1) (see also Chapter 7 and Bartels *et al.*, 2002).

Table 8.1.

Univariate Model fitting voor IQ at four ages (see also Bartels et al., 2002).

	MODEL	-2LL	df	$\Delta\chi^2$	Δdf	p	A[b]	C	E
IQ5	**ACE**	3180.40	411	2.20	3	.53	.26 *(.03-.52)*[a]	.50 *(.26-.68)*	.24 *(.18-.33)*
	AE	3193.71	412	13.31	1	.00			
IQ7	**ACE**	3054.37	378	3.04	3	.39	.39 *(.07-.72)*	.30 *(.00-.55)*	.31 *(.23-.44)*
	AE	3058.12	379	3.75	1	.05	.70 *(.60-.78)*	-	.30 *(.22-.40)*
IQ10	**ACE**	3140.87	388	5.39	3	.15	.54 *(.28-.83)*	.25 *(.00-.48)*	.21 *(.15-.29)*
	AE	3143.62	389	2.75	1	.10	.80 *(.72-.85)*	-	.20 *(.15-.28)*
IQ12	**ACE**	2908.72	377	5.01	3	.17	.64 *(.40-.88)*	.21 *(.00-.43)*	.15 *(.11-.22)*
	AE	2910.81	378	2.09	1	.15	.85 *(.79-.89)*	-	.15 *(.11-.21)*

[a] 95% confidence intervals; [b] A represents additive genetic influences, C represents shared environmental influences, and E represents nonshared environmental influences.

The aim of this study is to explore to what extent psychometric IQ and scholastic achievement, as assessed by the Dutch CITO-elementary test, are correlated and whether this correlation is due to genetic influences, shared or nonshared environmental influences common to CITO and intelligence or a combination of these influences. The unique aspect of this study is that the IQ data were collected longitudinally at ages 5, 7, 10, and 12 and that scholastic achievement is assessed at age 12. So we have the possibility to determine whether intelligence measured at age 5, 7, 10 and 12 may be used as a predictor of scholastic achievement at age 12. Since, scholastic achievement and IQ are also assessed at the same age a reliable measure of the association, without confounding effects related to age, can be obtained. Further, the variance found in the results of CITO will be disentangled into variance due to genetic influences, variance due to shared environmental influences (environmental influences shared by two members of a twin pair), and variance due to unique environmental influences (environmental influences unique to an individual). If shared environmental influences (C) determine the association, then we expects that IQ5-CITO will show the highest correlation, because C is of significant influence on IQ at age 5. If genetic factors (A) determine the association, then we expect the highest correlation between IQ12 and CITO, because genetic factors are the main source of individual differences of IQ at age 12.

Information on the strength of genetic and environmental influences on the results of the CITO and information on an association between CITO and intelligence at several ages is a valuable contribution to a discussion on the reliability of the CITO and use of this national test of educational achievement as a predictor of future scholastic success and the quality level of a certain elementary school in comparison to other elementary schools in the country.

METHODS

Participants

In 2000, the Netherlands Twin Register (NTR; Boomsma *et al.*, 1992; Boomsma, 1998; Boomsma *et al.*, 2002) started collecting the results of a national test of educational achievement (CITO) from all registered 12-year old twins. 85% of all Dutch schools yearly administer this test in the final class of elementary school. The main purpose of this test is to select for different levels of high school education. A standardized CITO score was collected for 1495 children, who took the CITO in 1998, 1999, 2000 or 2001.

Zygosity of this large CITO-sample was determined by DNA or blood group polymorphisms for 306 same-sex twin pairs. For the remaining same sex twin pairs zygosity was determined by discriminant analysis of questionnaire items. The questionnaire items allow accurate determination of zygosity of nearly 95%. The employment of the discriminant analysis and the use of zygosity questions are described in more detail in Rietveld *et al.* (2000b). In this sample of 1495 children there were 170 monozygotic female twin pairs (MZF), 113 dizygotic female twin pairs (DZF), 127 monozygotic male twin pairs (MZM), 113 dizygotic male twin pairs (DZM), and 168 dizygotic pairs of opposite sex twin pairs (DOS). There were 9 MZM incomplete twin pairs, 25 DZM incomplete twin pairs, 7 MZF incomplete twin pairs, 17 DZF incomplete twin pairs and 54 DOS incomplete twin pairs. For one child zygosity was missing.

A subsample of this NTR sample took part in a longitudinal study of the development of intelligence and problem behavior. This longitudinal study started in 1992 with recruitment of 209 twin pairs. The initial sample of 209 twin pairs was selected on the basis of age and zygosity of the twins, and their city of residence. Details on the demographic characteristics of the sample and information on parental occupation can be found in Rietveld *et al.*, (2000). Zygosity of the same-sex twins in this longitudinal sample was established by either blood group polymorphisms (137 pairs) or DNA analyses (24 pairs), and in a few pairs by physical resemblance assessed by the test-administrator (9 pairs). There were 47 monozygotic female (MZF), 37 dizygotic female (DZF), 42 monozygotic male (MZM), 44 dizygotic male (DZM), and 39 dizygotic pairs of opposite sex (DOS).

In this subsample IQ tests were administered at ages 5,7,10, and 12. Mean age at the first measurement occasion was 5.3 years (80% ranging from 5 years and 1 month to 5 years and 6 months). At the second measurement occasion mean age was 6.8 years (80% ranging from 6 years and 6 months to 7 years and 1 month). Mean age at the third measurement occasion was 10 years (80% ranging from 9 years and 11 months to 10 years and 1 month). Mean age at the fourth measurement occasion was 12 years (80% ranging from 11 years and 11 months to 12 years and 1 month).

The intelligence test was administered to all 209 twin pairs at age 5. At the second measurement occasion (age 7) 192 pairs of the original sample provided complete data on all subtests. The number of participating twin pairs increased to 197 when the children were tested around their 10th birthday. At the fourth measurement occasion (age 12) 192 twin pairs participated. A small group of four families refused consistently to participate after the first measurement occasion. Five families dropped out at both ages 10 and 12. The remaining nonparticipants refused participation at one measurement occasion. At age 5 and 12, one incomplete twin pair can be found in the data because of difficulties during testing (age 5) and refusal to participate (age 12). Due to serious loss of hearing one twin pair was assigned missing values at all four ages for IQ. This left a sample of 176 twin pairs with complete IQ data at all four ages. No significant difference in initial IQ (at age 5) was found for twins who dropped out on one or more occasions ($F_{3, 415} = 2.25$, $p = .082$). For at least 190 of the 209 twin pairs results for CITO at age 12 and Full-Scale IQ at age 5, or 7, or 10 or 12 are available.

Procedure

The Dutch CITO-elementary test

Educational achievement was assessed by the Dutch CITO-elementary test (Eindtoets Basisonderwijs 2002, Arnhem: Citogroep, 2002). The CITO consists of 240 multiple-choice items assessing four different intellectual skills: Language, Mathematics, Information Processing, and World Orientation. Each performance scale contains 60 multiple-choice questions. In 2001 the test slightly changed with respect to the distribution of the questions resulting in 60 questions for Mathematics and World Orientation, 90 questions for Language and 30 questions for Information Processing. Together the performance scales result in a standardized score between 501 and 550. The test is administered on three consecutive days in January or February when the children are in the final class of elementary school. In the present study the CITO data were collected by mail from the teacher, after informed consent from the parents or by mail from the parents as a question in a questionnaire on the child's behavior at age 12. In all analyses concerning the CITO score the mean is fixed to the population mean (534.5) in order to

control for volunteer bias, which in this respect could be a result from voluntary registration in the Netherlands Twin Register, voluntary sending in the results of the CITO, or voluntary participating in the CITO test (Neale and Cardon, 1992; Neale and Eaves, 1993). The population mean is based on a sample of 657869 children, who took the CITO in the years 1997 till 2001. The mean of our sample (m=537.88) is slightly, but significantly higher than this population mean (t_{1494} = 15.099, p=.00).

The intelligence tests

At ages 5 and 7 years the twins participated in a study on the development of cognitive abilities and brain-activity (Boomsma and Van Baal, 1998). At both measurement occasions the twins and their family visited the laboratory at the university. While one of the twins participated in the electro-physiological experiment, the co-twin participated in an intelligence test. At age 10 and 12 years a different procedure was followed. The twins and their parents could choose whether they preferred to come to the university or whether they preferred to be visited at home to participate in the intelligence test. The majority of the families (around 70% at both ages) preferred testing at home. No significant difference in IQ was observed between children tested at home or at the university (age 12 oldest of the twin: $F_{2,190}$=.654, p=.521; youngest of the twins: $F_{2.191}$=.32-12, p=.733). The intelligence test was assessed by an experienced test-administrator. All children received a present afterwards.

At age 5,7, and 10 the children were tested with the Revised Amsterdamse Kinder Intelligentie Test (RAKIT) (Bleichrodt *et al*, 1984). Six subtests, with age-appropriate items, were employed to assess cognitive functioning. The raw scores were standardized. For further details on this well-known Dutch intelligence test see Rietveld *et al.* (2000). At age 12 the twins completed the full version of the WISC-R. (Dutch version; Van Haasen *et al.*, 1986). The WISC-R consists of 12 subtests, 6 mainly verbal and 6 mainly non-verbal. The subtest scores are standardized, based on results of same-aged children in the Netherlands and the same standardization is used for boys and girls. Addition of the twelve standardized subtest scores results in Full-Scale IQ (IQ). The concurrent validity of the RAKIT and the WISC-R is .86 (Pijl *et al.*, 1984).

Descriptive Statistics

Descriptive statistics for full-scale IQ at age 5,7,10 and 12 (IQ5, IQ7, IQ10, and IQ12), and the standardized CITO scores were calculated using SPSS/windows 10. Differences in means and variances of IQ and CITO for boys and girls and monozygotic and dizygotic twins were tested with ANOVA. Twin correlations for the five zygosity groups (MZM, DZM, MZF, DZF, DOS) have been calculated to get a first impression of the genetic and

environmental influences on the variance in CITO scores. Pearson correlations were used to test the association (phenotypic correlation) between IQ at the four ages and CITO at age 12. MZ and DZ cross-correlations are calculated to get an impression of influences of genes and environment on the covariance between IQ and CITO.

Genetic Modeling

Genetic model fitting of twin data allows for separation of the observed phenotypic variance into its genetic and environmental components. Additive genetic variance (A) is the variance that results from the additive effects of alleles at each contributing genetic locus. Shared environmental variance (C) is the variance that results from environmental events common to both members of a twin pair. Unique environmental variance (E) is the variance that results from environmental effects that are not shared by members of a twin pair. Estimates of the unique environmental effects also include measurement error. To account for this source of variance, E is always specified in the model.

The different degree of genetic relatedness between monozygotic (MZ) and dizygotic (DZ) twin pairs was used to estimate the contribution of these factors to the phenotypic variation in IQ at the four ages and in CITO scores. Similarities for MZ twins are assumed to be due to additive genetic influences plus environmental influences that are shared by both members of a twin pair. Experiences that make MZ twins different from one another are unique environmental influences. Because DZ twins share 50% of their genetic material on average, like other siblings, genetic factors contribute only half to their resemblance. As for MZ twins the shared environment contributes fully. Model fitting to twin data is based on the comparison of the variance-covariance matrices in MZ and DZ twins. Exploiting the known difference in genetic contribution to intra-pair resemblance of MZ and DZ twin pairs, influences of additive genetic, shared environmental and unique environmental factors are estimated using the computer program Mx (Neale *et al.*, 1999).

Univariate model fitting was carried out to estimate the genetic and environmental components in CITO scores. Per time point (CITO-IQ5, CITO-IQ7, CITO-IQ10 and CITO-IQ12) a bivariate model (Cholesky decomposition) was used to estimate genetic and environmental influences (Figure 8.1). Rather than decomposing the variance of IQ and CITO into genetic and environmental sources of variance, bivariate genetic analysis decomposes the variance of each measured variable and the covariance between the measured variables into genetic and environmental sources.

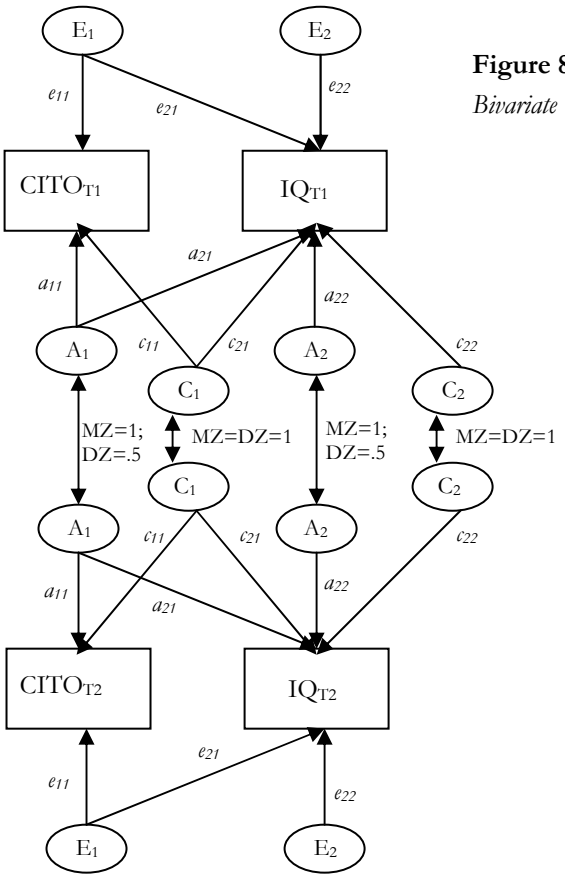


Figure 8.1.
Bivariate Cholesky Decomposition

To make optimal use of all available data, the analyses were performed on raw data. Submodels were compared by hierarchic χ^2 tests. The χ^2 statistic is computed by subtracting –2LL for a reduced model from that for the full model. This resulted in a χ^2 statistic ($\chi^2 = 2LL_0 – (-2LL_1)$). We compared alternative models by means of the principle of parsimony. We began the bivariate model fitting with a full model with additive genetic, shared environmental and unique environmental influences (ACE model), including sex-differences in mean and a free estimate of the degree of genetic relatedness in twins of opposite sex (model 1). First, we tested whether different genes influence IQ and CITO in boys and girls (model 2). It was also tested whether the influences of the genes are of different magnitudes in boys and girls (model 3). Further we tested whether the influences of unique environment were specific for CITO and IQ (model 4). After these model reductions it was tested whether the covariance between IQ and CITO is based on a common genetic background, a common source of shared environmental influences or

both. Estimates of genetic, shared environmental and unique environmental influences on CITO and the covariance between IQ (four ages) and CITO have been estimated based on the best fitting model.

RESULTS

Descriptive statistics for IQ at the four ages and CITO at age 12 are presented in Table 8.2. No differences in means are found for boys and girls or monozygotic and dizygotic twins for IQ and CITO. Significant differences in variances between boys and girls are only found for IQ at age 10($F_{1,390}$=4.326, p=.038). No differences in variances are found for monozygotic and dizygotic twins.

Table 8.2.

Descriptive statistics for Full-Scale IQ at different ages and CITO

		N	Min	Max	Mean	SD	Skewness		Kurtosis	
								s.e.		s.e.
IQ5	♂	210	64	139	102.32	13.19	-.164	.168	.281	.334
	♀	205	70	142	103.20	13.19	.047	.170	.142	.338
IQ7	♂	194	65	139	102.98	14.63	-.079	.175	.001	.347
	♀	188	62	145	102.80	14.75	-.176	.177	.076	.353
IQ10	♂	195	69	145	107.75	14.40	-.009	.175	.040	.346
	♀	197	63	145	106.17	16.59	-.067	.173	-.361	.345
IQ12	♂	185	66	138	101.03	13.00	.123	.179	.216	.355
	♀	196	61	127	99.08	13.32	-.171	.174	-.018	.346
CITO12	♂	702	510	550	538.13	8.61	-.865	.092	.317	.184
	♀	793	506	550	537.66	8.70	-.730	.087	.073	.173

Phenotypic correlations between IQ and CITO are presented in Table 8.4 (upper part). All correlations are significant at the α= .01 level, indicating medium to strong associations between IQ at several ages and CITO. The expected rise in correlation from IQ5 and CITO to IQ12 and CITO is observable in the final column of the table. This rise is not surprising because of the fact that the CITO is taken at age 12 and is in that sense most comparable to IQ12. Twin correlations for the five zygosity groups for CITO and the univariate model fitting results are presented in Table 8.3.

Fifty-seven percent of the individual differences in CITO can be explained by additive genetic influences. Shared environmental influences explain 27% and nonshared

environmental influences explain 16% of the total variance respectively. Univariate model fitting showed no presence of sex-differences in heritability for CITO.

Twin cross correlations for monozygotic (MZ) and dizygotic (DZ) twins were calculated separately to explore the genetic and environmental influences on the observed association between CITO and IQ. As can be seen in Table 8.4 (lower part), the MZ cross correlations over time (above the diagonal) are higher than the DZ cross correlations over time (below the diagonal), suggesting that the observed significant association between CITO and IQ at four ages is at least partly due to genetic factors. The MZ correlations, though, are not twice as high as the DZ correlation which indicates influences of shared environment as well. Further, when the correlations of the adjoining age-intervals are compared (CITO-IQ5; CITO-IQ7; CITO-IQ10; CITO-IQ12) the increased difference between MZ and DZ correlations, from age 5 to age 10, suggests an increase in the genetic contribution to the association in this age interval.

Table 8.3.

Twin correlations and Univariate model fitting results for CITO at age 12.

	MZF[a]		**DZF**		**MZM**		**DZM**	**DOS**
CITO	.85 *(.80-.89)[b]*		.47 *(.30-.61)*		.83 *(.77-.88)*		.56 *(.42-.67)*	.55 *(.44-.65)*
Model	**-2LL**	**df**	**$\Delta\chi^2$**	**Δdf**	**p**	**A[d]**	**C**	**E**
ACE+sd[c]	10410.65	1491						
ACE	10411.11	1494	.46	3	.93	.57 (.44-.71)	.27 (.13-.39)	.16 (.13-.19)
AE	10424.19	1495	13.8	1	.00			
CE	10494.41	1495	83.3	1	.00			

[a] MZF= monozygotic female, DZF= dizygotic female, MZM= monozygotic male, DZM=dizygotic males, DOS= dizygotic opposite sex; [b] 95% confidence intervals; [c] model with sex-differences in the strength of the additive genetic, shared environmental and nonshared environmental influences; [d] A represents additive genetic influences, C represents shared environmental influences, and E represents nonshared environmental influences.

Model fitting results of the bivariate Cholesky Decomposition for CITO with IQ at the four ages are presented in Table 8.5. As expected from the univariate model-fitting procedure, no sex-differences are found (models 2 and 3). Further, the estimated genetic and shared environmental correlations indicate overlapping influences for genetic and shared environmental effects on CITO and IQ. For CITO-IQ5, CITO-IQ7, and CITO-IQ10 the unique environmental influences could be reduced to the variable specific influence only (model 4). For CITO-IQ12, however, some overlap in nonshared

environmental influences is observed. For CITO-IQ5 and CITO-IQ7 no difference was observed between a model with a common factor for additive genetic influences or a model with a common factor for shared environmental influences. Both models did not significantly worsen the fit. However, model reduction to a common factor for both additive genetic and shared environmental influences did change the χ^2 significantly. For CITO-IQ10 both the additive genetic and the shared environmental influences could be reduced to a common factor influencing CITO and IQ10. For CITO-IQ12, the full model could be reduced to a model with a common factor for shared environmental influences.

Table 8.4.

Phenotypic cross correlations for IQ at four ages and CITO, calculated for the complete dataset (upper part table) and MZ (above diagonal) and DZ (below diagonal) cross correlations (lower part table).

Total Sample	**IQ5**	**IQ7**	**IQ10**	**IQ12**	**CITO**
IQ5	1.00	.65 *(.59-.70)*	.65 *(.59-.70)*	.64 *(.57-.69)*	.41 *(.31-.50)*
IQ7		1.00	.72 *(.67-.77)*	.69 *(.63-.74)*	.50 *(.40-.58)*
IQ10			1.00	.78 *(.74-.82)*	.60 *(.52-.66)*
IQ12				1.00	.63 *(.55-.69)*
CITO					1.00
DZ\MZ	**IQ5b[b]**	**IQ7b**	**IQ10b**	**IQ12**	**CITOb**
IQ5a[a]	-	.66 *(.54-.75)*	.67 *(.56-.76)*	.68 *(.57-.77)*	.35 *(.12-.53)*
IQ7a	.42 *(.30-.54)*	-	.71 *(.60-.79)*	.68 *(.57-.77)*	.47 *(.27-.62)*
IQ10a	.42 *(.30-.54)*	.39 *(.26-.52)*	-	.79 *(.71-.85)*	.55 *(.36-.68)*
IQ12a	.42 *(.29-.54)*	.42 *(.29-.54)*	.45 *(.32-.57)*	-	.52 *(.32-.65)*
CITOa	.37 *(.23-.48)*	.37 *(.24-.48)*	.37 *(.24-.48)*	.50 *(.39-.59)*	

[a] a is the oldest of the same sex twin pair and the boy in twin pairs of opposite sex; [b] b is the youngest of the same sex twin pair and the girl in twin pairs of opposite sex

The percentage of variances explained by additive genetic, shared environmental and unique environmental influences based on the full model (ACE without sex-differences) are presented in Table 8.6. From age 5 to age 10, the expected increase in heritability of IQ can be seen. This increase was previously found in these data (Bartels *et al.*, 2002).

TABLE 8.5.
Bivariate model-fitting results for CITO-IQ

	MODEL	-2LL	df	χ^2	df	c.t.m.[a]	p
CITO-IQ5	**Model 1.** ACE, H_{free}, sex differences	13540.560	1883				
	Model 2. ACE sex differences	13540.576	1884	.016	1	1	.90
	Model 3. ACE no sex differences	13547.738	1893	7.162	9	2	.62
	Model 4. ACE **E**:variable-specific factors only	13548.158	1894	.42	1	3	.52
CITO-IQ7	**Model 1.** ACE, H_{free}, sex differences	13412.527	1850				
	Model 2. ACE sex differences	13412.546	1851	.019	1	1	.89
	Model 3. ACE no sex differences	13417.961	1860	5.415	9	2	.78
	Model 4. ACE **E**:variable-specific factors only	13417.962	1861	.001	1	3	.92
CITO-IQ10	**Model 1.** ACE, H_{free}, sex differences	13447.831	1860				
	Model 2. ACE sex differences parameter estimates	13447.837	1861	.006	1	1	.94
	Model 3. ACE no sex differences	13462.791	1870	14.954	9	2	.09
	Model 4. ACE **E**:variable-specific factors only	13462.798	1871	.007	1	3	.93
CITO-IQ12	**Model 1.** ACE, H_{free}, sex differences	13210.245	1849				
	Model 2. ACE sex differences parameter estimates	1321.343	1850	.01	1	1	.92
	Model 3. ACE no sex differences	13220.468	1859	10.125	9	2	.34
	Model 4. ACE **E**:variable-specific factors only	13225.776	1860	5.308	1	3	.00

[a] c.t.m. = compared to model

Further, additive genetic effects explain around 60% of the individual differences in CITO. 24% of the variance in CITO can be explained by shared environmental influences,

while unique environmental influences explain 16%. This pattern of influences is identical to the results of the univariate analysis. Covariance between CITO and IQ at the four ages could be mainly explained by additive genetic factor for CITO-IQ7 and CITO-IQ10 (Table 8.6). The covariance between CITO-IQ5 and CITO-IQ12 was accounted for by additive genetic factors as well as shared environmental factors. As for the most important covariance, CITO-IQ12, it is indicated that the same shared environmental influences influence both CITO and IQ at that age. Some nonshared environmental influences on the covariance between CITO and IQ 12 are observed. These influences suggest idiosyncratic experience specific for age 12.

In summary, additive genetic as well as shared environmental effects are of significant influence on the observed association between CITO and IQ at four ages. Heritabilities of IQ rise from age 5 to age 10 and the heritability of CITO is around 60%.

DISCUSSION

Previous studies indicate that in addition to several environmental variables, intelligence seems to be an explaining factor for the variance in scholastic achievement (Eaves and Darch, 1990, Jensen, 1972). Behavior genetic studies also established that genetic effects are the primary source of variance underlying the observed association between cognition and achievement (Petrill *et al.*, 1993; Wadsworth *et al.*, 1995a, 1995b, Thompson *et al.*, 1991).

The primary aim of this study was to establish if an significant association exists between cognitive abilities and a national test of educational achievement (CITO) in a Dutch sample and to establish the background mechanism of this possible phenotypic correlation. Further the predictive value of IQ for scholastic achievement was examined. To this end multivariate behavior genetic analyses with CITO and IQ at age 5, 7, 10 and 12 were conducted. The results point to genetic effects as the main source of variance in CITO score and an important source of covariance between CITO score and IQ. Beside genetic influences, shared environment shows significant influences on the variance of CITO and IQ and the covariances at all ages. Further, based on correlation between IQ5 and CITO and taking the results of the bivariate model fitting into account IQ5 seems to be an accurate indicator for CITO at age 12. However, the association is not strong enough to completely predict outcomes of the CITO at age 12 from cognitive ability at age 5. Despite the wealth of evidence for small but significant sex-differences in cognitive abilities (for a review see Helgeson, 2002) no sex-differences for CITO or IQ were found. CITO scores were available for 702 boys and 793 girls, resulting in means of 538.13 and 537.66 respectively.

Table 8.6.

Percentage of variance and covariance explained by A, C en E, based on a Cholesky ecomposition without sex-differences, with 95% confidence intervals

	Variance	A[a]	C	E
Model 4	**CITO**	.57 *(.44-.72)*	.27 *(.13-.39)*	.16 *(.13-.19)*
	IQ5	.22 *(.01-.46)*	.55 *(.32-.72)*	.23 *(.17-.32)*
Model 4	**CITO**	.56 *(.44-.70)*	.28 *(.14-.40)*	.16 *(.13-.19)*
	IQ7	.40 *(.11-.73)*	.29 *(.00-.54)*	.31 *(.22-.43)*
Model 4	**CITO**	.58 *(.45-.72)*	.26 *(.13-.39)*	.16 *(.13-.19)*
	IQ10	.50 *(.26-.79)*	.30 *(.02-.52)*	.20 *(.15-.28)*
Model 3	**CITO**	.55 *(.43-.68)*	.29 *(.16-.40)*	.16 *(.13-.19)*
	IQ12	.51 *(.31-.73)*	.33 *(.13-.52)*	.15 *(.11-.22)*
	Covariance	**A**	**C**	**E**
Model 4	**CITO-IQ5**	.40 *(.00-.92)*	.60 *(.08-1.00)*	-
Model 4	**CITO-IQ7**	.75 *(.32-1.00)*	.25 *(.00-.68)*	-
Model 4	**CITO-IQ10**	.83 *(.53-1.00)*	.17 *(.00-.47)*	-
Model 3	**CITO-IQ12**	.41 *(.19-.64)*	.51 *(.28-.70)*	.09 *(.01-.16)*
		A correlation	**C correlation**	**E correlation**
Model 4	**CITO-IQ5**	.42	.58	-
Model 4	**CITO-IQ7**	.74	.42	-
Model 4	**CITO-IQ10**	.90	.35	-
Model 3	**CITO-IQ12**	.47	1.00	.47

[a] A represents additive genetic influences, C represents shared environmental influences, and E represents nonshared environmental influences.

In comparing the estimates of genetic and environmental influences on IQ in the univariate analyses to the estimates in the bivariate analyses (with CITO), differences are observed. Based on the bivariate analyses, the estimates of additive genetic influences are slightly lower and the estimates of the shared environmental influences are slightly higher. This difference can be explained by the increase of power by multivariate instead of univariate analyses. Multivariate analysis increases the power to detect shared environmental influences, especially when the bivariate analyses are conducted with a trait on which shared environmental influences are significant (e.g. CITO). The correlation for shared environmental influences on CITO at age 12 and IQ at age 12 is 1, which indicates

that the same shared environmental factor influences CITO and IQ. This overlap in shared environmental influences results in an increased power to detect these influences in a bivariate analysis.

Remarkable is the drop in genetic influences on the covariance between CITO and IQ at age 12. An possible explanation can is suggested based on the results of the longitudinal study previous conducted in this sample (Bartels *et al.*, 2002). In this longitudinal study a common genetic factor is found, which influences cognitive ability at all ages. So it can be hypothesized that the genes that influences stability in IQ also influence the covariance between CITO and IQ at ages 5, 7, 10 and 12. The longitudinal study in cognitive abilities shows that shared environmental influences are partly explained by a common factor and partly by age specific factors. The covariance between CITO score and IQ at age 12 however, can be based on several time specific influences. For instance age specific shared environmental influences explain a large part of the covariance between CITO and IQ at age 12. So the fact that CITO is measured at age 12 and analyzed in combination with IQ assessed at age 12, makes this bivariate analyses more sensitive than the bivariate analyses for CITO and IQ at previous ages.

Focusing on genetic influences as the overlapping factor for the association, the large CITO database creates opportunities for future research on the genetics of cognition. Administering an intelligence test is time consuming and in order to get more insight in the genetic background of cognition large sample sizes are necessary. Especially since the CITO is a nationwide standardized test, the use of the database and the possibilities to recruit parents, siblings and normal controls for genetic studies would boost power to finally find genes influencing cognitive abilities.

Focussing on shared environmental influences as the main overlapping factor for the association between CITO and IQ, it is interesting to focus on the exact nature of these environmental influences. In general, family environment (SES) is considered to be the main factor of shared environmental influences. However, studies nowadays, also emphasize aspects outside the family environment, like friends or being a member of a sportsclub, which may also cause similarities between two children of a twin pair. Obviously in measuring scholastic achievement and cognitive abilities and taking the Dutch school system into account, one may also consider the school environment as an important source of shared environmental influences. Information on 'same or different' teachers indicated that out of a large sample of 12-year-old twin pairs (n=1164) 63% of twins are taught by the same teacher, whereas 37% go to separate classes. This ratio makes teacher or classroom environment a shared environmental influence for the majority of the children. Since, in the Dutch school system children move to a different teacher each school year, this results in a lack of continuity in this particular aspect of shared

environment. So, these shared but age-specific experiences within the classroom may be represented by the age-specific factors as specified significant in a previous longitudinal study (Bartels *et al.*, 2002). Further indication to consider the classroom and teacher as shared environment is given by preliminary results on twin correlations for CITO in the same large sample of 12- year- old-twin pairs. The twin correlation for CITO in twins taught by the same teacher indicate higher influences of shared environment than the twin correlations for CITO in twins taught by different teacher. It should be noted that because only a minority of the twin go to separate classes the zygosity groups to calculate these twin correlation are still small. The collection of CITO data and data on 'different or same' teacher is a continuous process at the NTR, so more insight in this matter can be gained in the future. The unique environment was found to explain only a small portion of the variance of CITO It further seems to be of no influence on the association between CITO and IQ, except for the association between CITO and IQ12. The finding of the influence of nonshared environmental influences on this covariance indicates that, besides measurement error, pure idiosyncratic experience are of importance for individual differences in cognitive abilities and CITO at age 12. Further, the finding of a significant influences on the association between CITO and IQ at age 12 only, underlines the transient nature of these idiosyncratic experiences. This transient nature of nonshared environmental influences was previous found in a longitudinal study on the development of intelligence (Bartels *et al.*, 2002)

Cognitive abilities seem to be an explaining factor for the variance in scholastic achievement as measured with the CITO. The association between CITO and IQ is both mediated by underlying genetic and shared environmental influences. Further it is clear that genetic background accounts for almost 60% of the individual differences found in CITO scores in a large sample of 12-year-olds. The large heritability indicates that the CITO is a valuable instrument to measure capacities in cognition in children but may not be the correct instrument to put the effect of education in a specific school to the test.

With the unique databases of CITO in mind, future studies could be very valuable. For instance, influences of classroom and teacher as a source of shared environmental influences could be sorted out by comparing twins attending the same class with twins attending separate classes. Further, with the value put on the results of the CITO nowadays, it is important to sort out the background of the individual differences in the test results. Another valuable future study could focus on comparison of the decomposition of the variance in CITO scores measured in different cohorts. Daily newspapers mostly devote their articles to the reliability of the national test of educational achievement. Questions are raised about the measurement procedure and the non-standardized preparation of the children. There is no control on the amount of practice

prior to the actual days of testing. Opponents of the CITO often use these arguments in discussing the value of the test. The high correlation between MZ twins, which can be regarded as an alternative measure of test-retest reliability suggests that reliability of the CITO test is good to excellent.

In the USA it has been proven that supplementing teachers opinions with standardized screening test results is needed to ensure accurate decision-making (Glascoe, 2001). The teacher's opinion may be biased by formal expectations (Demaray and Elliot, 1998), knowledge of the child's SES (Lichtenstein, 1984), or risk factors for difficulties such as language spoken at home (Glascoe, 2001). These previous studies emphasize the importance of independent testing in order to advice parents on future educational level of their children. The current study on CITO and IQ at four ages may give rise to a valuable discussion on the reliability and appropriateness of the CITO as a measure for the quality level of a certain elementary school in comparison to other elementary schools in the country.

Limitations of the study

It is know that academic achievement or educational attainment is a phenotype on which nonrandom mating occurs. Besides the so-called passive elements of mate selection (e.g. type and length of education, social class, area of residence), active personal preferences for physical and psychological attributes, including IQ, may play a role in mate selection. This will induce positive assortative mating. Assortative mating is important for genetic research for two reasons. First, assortative mating increases genetic variance in a population. In other words, positive assortative mating increases variance in that the offspring differ more from the average than they would if mating were random. Even if spouse correlations are modest, assortative mating can greatly increase genetic variability in a population, because its effects accumulate generation after generation. Assortative mating is also important because it affects estimates of heritability. Positive assortative mating increases the resemblance between dizygotic twins because it renders the parents of these twins more similar compared to the situation where assortative mating is absent. Identical or monozygotic twins, however, are already at the point of maximum genetic resemblance, and are thus unaffected by positive assortative mating (Plomin *et al.*, 2000). Based on the comparison of MZ correlations and DZ correlations, an increased DZ correlation will result in decreased estimates of heritability, when estimating C. It is possible that in our study the estimate of shared environmental influences is inflated by assortative mating.

Strong assortment effects have been shown for cognitive abilities (Mascie-Taylor, 1989; Nagoshi *et al.*, 1987; Philips *et al.*,1988; Tambs *et al.*, 1989). The extension of the

classical twin design to include parental measurements can be used to disentangle sources of variation that are confounded in the classical twin design and explicitly assess the roles of mate selection in the determination of scholastic achievement (Eaves *et al.*, 1989). In future research, information on CITO scores for the parents of the twins should be collected from the unique database of CITO to sort out the presence and strength of assortative mating in these traits.

9

Heritability of Cortisol Levels: Review and Simultaneous Analysis of Twin Studies

This chapter is in press as:

Bartels, M., Van den Berg, M, Sluyter, F., Boomsma, D.I., and de Geus, E.J.C. (2002). Heritability of Cortisol Levels: Review and Simultaneous Analysis of Twin Studies. *Psychoneuroendocrinology*, in Press.

ABSTRACT

Cortisol has a pivotal role in physical and mental health, but relatively few studies have paid attention to individual differences in cortisol levels and the etiology of these differences, in particular their possible genetic basis. In this article we review the existing literature on the heritability of cortisol levels. Most of the studies, which have been carried out in genetically informative samples, lack methodological consistency with regard to frequency and timing of sample collection. The circadian rhythm in cortisol levels was often not taken into account. A power analysis shows that none of these studies used adequate sample sizes to distinguish genetic from shared environmental influences as a cause for familial aggregation. Results of a simultaneous analysis of 5 comparable twin studies suggest a heritability of 62%. Hence, we conclude that, to understand the contribution of genetic and (shared) environmental influences to variation in basal cortisol levels, future studies should be designed more rigorously with strict collection and sampling protocols, sufficient sample size and repeated measures across multiple days.

INTRODUCTION

Cortisol is a steroid hormone secreted by the outer cortex of the adrenal gland. Its secretion is stimulated by ACTH (adrenocorticotrophic hormone), produced in the pituitary in response to corticotropin-releasing hormone (CRH), a product from the neurons in the paraventricular nucleus of the hypothalamus. In the characteristic diurnal rhythm of plasma cortisol level typically 10-15 well-defined pulses of variable amplitude are observed, with a morning maximum, declining levels throughout the daytime, a period of low concentrations generally centered around midnight, and an abrupt rise after the first few hours of sleep (Weitzman, 1971). This diurnal cycle is tied to the sleep-wake cycle and to the light-dark cycle (Spith-Schwalbe *et al.*, 1993). Although both mechanisms are involved in the regulation of the HPA axis, the light-dark cycle is still the primary synchronizer of a basically endogenous rhythm originating in the suprachiasmatic nucleus of the hypothalamus (Fischman *et al.*, 1988; Van Cauter *et al.*, 1990; Boivin *et al.*, 1996; Scheer and Buijs, 1999). The circadian rhythm is reversed in nocturnal species and disrupted by sleep deprivation, and changes in the sleep pattern (e.g. shift work and long distance travel). Further, plasma cortisol release is tightly regulated through negative feedback at the pituitary, hypothalamus and hippocampus (Kovacs *et al.*, 1987; Jacobson and Sapolsky, 1991). Strength of this feedback signal strongly varies with time of day (Dorin *et al.*, 1996; Huizinga *et al.*, 1998; Young *et al.*, 1998), contributing to the characteristic diurnal rhythm in plasma cortisol levels. After its release, the major proportion of cortisol binds to the plasma proteins corticosteroid binding globulin (CBG or transcortin) and albumin, which prevents the hormone from penetrating the membranes of their target cells. Only, about 3-5% of the total cortisol is the unbound, biologically active fraction.

Large individual differences exist in basal cortisol levels at all points of the circadian cortisol curve (Smyth *et al.*, 1997). These individual differences in cortisol levels play a prominent role as an explanatory variable in studies on physical (Walker, 1996; Mantero and Boscaro, 1992; Pedersen and Hoffman-Goetz, 2000; Roy *et al.*, 2001; Rosmond and Bjorntorp, 2000c;) and mental health (Young *et al.*, 2000; Posener *et al.*, 2000; Goodyer *et al.*, 2000). There are many sources of individual differences in cortisol levels, including negative feedback regulation through the corticoid receptors. In this regulation two receptor types can be distinguished: the mineralocorticoid (MR, or type-I) receptor and the glucocorticoid (GR, or type-II) receptor (Veldhuis *et al.*, 1982; Reul and De Kloet, 1985). Because of its much higher affinity to cortisol, MRs are predominantly occupied under (nonstress) basal levels whereas during stress, when cortisol levels are much higher, GRs become increasingly occupied (Young *et al.*, 1998). Individual differences in the number, affinity and efficiency of signaling cascades of these receptors will directly affect

cortisol levels and biological activity. Further, individual differences may arise from the secretion of ACTH in response to CRH or the secretion of cortisol by the adrenal cortex in response to ACTH (Dorin *et al.*, 1996; Posener *et al.*, 1997; Beuschlein *et al.*, 2001). Finally, basal cortisol levels are responsive to individual differences in the capacity of 11β-Hydroxy steroid dehydrogenase (11β-HSD), that causes the conversion of the biologically active cortisol to its inactive metabolite cortisone.

Ultimately, individual differences in all these mechanisms arise from two main factors: genetic and environmental influences. As for the latter, evidence suggests that early adverse experiences, like childhood abuse or parental separation, play a prominent role in development of mood and anxiety disorders and that corticotrophin-releasing hormone (CRH) systems may mediate this association (Mullen *et al.*, 1996; Heim *et al.*, 2000). Further evidence for this association has been assembled in animal models, where prenatal and early developmental stress, often related to parental rearing, have been shown to cause long-lasting or even permanent alteration of the HPA axis (Plotsky and Meaney, 1993; Levine, 1994). Not only early experiences, but also experiences later in life can influence HPA axis activity. For example, trauma survivors with posttraumatic stress disorder such as Vietnam veterans, holocaust survivors or victims of abuse are characterized by decreased urinary cortisol level as compared to healthy controls (see, among others, Yehuda et al, 1991; 1995, 2000). Accordingly, environmental challenges are important in the development of HPA axis disregulation and stress-related diseases. Still, this does not answer the question why similar stressors affect some individuals strongly, while others remain relatively untouched. These remaining individual differences point in the direction of genetic influences on variation in cortisol levels.

A powerful tool to unravel the genetic architecture of individual differences is to study genetically related individuals. Family studies might give a first impression of familial aggregation, but they can not distinguish between genetic and shared environmental effects. Similarities between family members may be created either by genetic relatedness or by sharing the same family environment, the so-called shared environment (C). A method that solves this problem, is the classical twin design. Monozygotic (MZ) twins derive from a single zygote and therefore two individuals of a MZ twin pair are genetically identical. Dizygotic (DZ) twins develop from two distinct zygotes and share on average 50% of their genes, like 'ordinary' brothers and sisters. Hence, the only possible way to explain the variation in cortisol levels between two members of a MZ twin pair are environmental effects that are not shared by those two: the so-called nonshared environmental influences (E). Conversely, the variation in cortisol levels between two members of a DZ twin pair could result from different genes and/or nonshared environmental influences. Accordingly, the difference in relatedness between MZ and DZ

twin pairs (mostly expressed as correlation coefficients: r_{MZ} and r_{DZ}) gives information about the strength of the genetic and environmental influences on the trait under investigation (Martin *et al.*, 1997). It further allows the separation of environmental influences into those of the environment shared by members of a family and those unique for each individual.

Twin and family studies constitute a powerful instrument, but surprisingly few attempts have been made to estimate the impact of genetic and environmental factors on the regulation of cortisol levels. The first and main purpose of this article was to review the existing studies, listed in table I, to obtain insight in the genetic and environmental influences on cortisol levels. Using PubMed and the search terms *twin, cortisol, corticosteroid, heritability* and *family*, 29 studies emerged. However, the studies by Norman *et al.* (1982; 1983a; 1983b; 1984) and Lopez Bernal *et al.* (1980) are based on neonates or twin pregnancies, which is beyond the scope of this paper. Further, several case studies (Mendlewicz *et al.*, 1984; Milford *et al.*, 1994; Li *et al.*, 1997; Pinheiro *et al.*, 1999) have been omitted, since no reliable heritability can be estimated based on one case. Likewise seven studies (Nurnberger *et al.*, 1983; Linkowski *et al.*, 1985; Schuckit *et al.*, 1991; Karl and Schulte, 1994; Smyth *et al.*, 1997; Walker *et al.*, 1998; Yehuda *et al.*, 2000) have been excluded because these are based on unrelated subjects or are family studies without any information on parent-offspring correlations, which gives no opportunity to estimate genetic parameters. A study by Schwartz *et al.* (1972) has been excluded because it is based on an ocular response to dexamethasone eye-drops. Finally, one study was published in Polish (Raczynska *et al.*, 1978).

What immediately catches the eye in Table 9.1 is the huge variation in heritability estimates (0.0% to 84%, with a median of 52%). To explain this discrepancy, a secondary purpose of this paper was to scrutinize the methodological aspects of existing studies to select studies with comparable methodology for a simultaneous (or meta-) analysis of the MZ and DZ correlations reported in these studies. Three fundamental issues were addressed: how the samples were collected, when they were collected, and how they were analyzed in the laboratory.

Collection methods

Table 9.1 shows that different methods of collection –blood, saliva, urine - have been used over the years. In general, saliva collection is the most practical and stress-free method of cortisol collection in a large group of subjects (both adults and children). The

Table 9.1. *Overview of genetic studies conducted in a twin design*

Study	*Subjects*	*Cortisol measurements*	*Results / direction of association*
Saliva cortisol sampling			
Young *et al.* (2000)	- 29 MZF[a]	- within 45 minutes after awakening - immediately before bedtime - during a two-week period - DPC Coat-account Cortisol assay kits	- 40-45% of total variance in salivary cortisol shared by MZ twins
Wüst *et al.* (2000)	- 31 MZF - 21 MZM - 13 DZF - 13 DZM - 26 DOS	- 0, 30, 45, and 60 minutes after awakening - at 0800h, 1500h and 2000h - over a two-day period - RIA	- significant genetic influence on the cortisol response to awakening - day-time cortisol profile is under influence of shared environmental factors - h^2 (mean increase) = .40 [b] - h^2 (area under the curve)= .48
Kirschbaum *et al.* (1992)	- 7 MZF - 6 MZM - 11 DZ (12 F, 10 M)	- baseline - after ergometric exercise (ERG) - after psychological stress (PSY) - after CRH stimulation (CHR) - Between 1700-1830h - at a 10 min interval from –35 min (CRH) or –10 min (ERG, PSY) - through 90 minutes after stimulation onset - RIA	- genetic factors are involved in adrenocortical response to CRH stimulation, psychological stress and for the baseline condition, but not for physical exercise. - Strong intraindividual stability was found for females - MZ intraclass coefficients significant higher than DZ intraclass coefficients for baseline CRH, ERG and PSY and peak CRH - h^2 (baseline CRH) = .72 - h^2 (baseline ERG) = dominance - h^2 (baseline PSY) = .52 - h^2 (peak CRH) = .84
Plasma cortisol sampling			
Froelich *et al.* (2000)	- 23 MZM - 28 MZF - 16 DZM - 21 DZF	- baseline at 0750-0800h - 15, 60, 75, 120, 180, and 240 minutes after onset of drinking - RIA	- cortisol responses to alcohol showed no signs of heritability - basal cortisol levels (2 hours before drinking) showed no signs of heritability
Pritchard *et al.* (1999)	- 7 MZM	- between 0730-0800h - after cycle-ergometer exercise - RIA	- observed changes in cortisol levels were correlated for both twins - no significant twin resemblance for basal plasma cortisol levels
Inglis *et al.* (1999)	- 75 MZ - 71 DZ	- 09.00h, after 30 minutes of rest - 30 minutes after ACTH stimulation - RIA	- h^2 (basal plasma cortisol) = .46 - cortisol response to ACTH stimulation showed no signs of heritability
Pritchard *et al.* (1998)	- 12 MZM	- baseline levels and 24 hours after termination of overfeeding between 0730-8.00h after overnight fasting - RIA	- no genetic control of cortisol levels found in response to overfeeding

Table 9.1. – continued *Overview of genetic studies conducted in a twin design*

Study	*Subjects*	*Cortisol measurements*	*Results / direction of association*
Linkowski *et al.* (1993)	- 11 MZM - 10 DZM	- every 15 minutes over a 25-hour period. - first hour is not considered - RIA	- genetic factors control absolute value of morning acrophase, timing nocturnal nadir and to some extent proportion of pulse variation - environmental factors contribute to morning acrophase - h^2 (pulsatility) = .74[a] - h^2 (morning acrophase) = .68 - h^2 (timing nocturnal nadir) = .48
Meikle *et al.* (1988)	- 20 MZM - 20 DZM	- 3 samples between 8 am and 9.30 am at 20 min. interval - levels of bound and unbound cortisol and DHEA-S - RIA	- heredity and environmental factors influence total bound and unbound cortisol levels - h^2 (total cortisol) = .51 - h^2 (unbound) = .58
Nurnberger *et al.* (1982)	- 13 MZ - 3 DZ - 11 patients with bipolar affective disorder	- –35, -20, -5, +5, +15, +30, +45, +60 and +90 minutes and +3, +6, and +24 hours after amphetamine stimulation - timing stimulation between 0900h-10.45h - RIA	- no significant twin concordance in basal cortisol - no significant twin concordance in change in cortisol after amphetamine stimulation
Maxwell *et al.* (1969)	- 66 MZ - 76 DZ	- plasma 11-hydroxycorticosteroids were measured after determination of BP - both twins at the same time - fluorimetric method	- genetic factors contribute to variation of corticosteroid levels in females, not in males
Urine cortisol sampling			
Inglis *et al.* (1999)	- 75 MZ - 71 DZ	- collected in a 24-hour period. - corticosteroid metabolite excretion was measured - gas chromatography-mass spectrometry	- strong heritability for levels of cortisol metabolites in urine is found - h^2 (24-h tetrahydrodeoxycortisol) = .59

[a] MZM = monozygotic male, DZM= dizygotic male, MZF= monozygotic female, DZF= dizygotic female, DOS= dizygotic opposite sex
[b] heritability (h^2) calculated on given MZ and DZ correlations

reason why blood and urine sampling have been used more often, is probably historical as the development of the 'Salivette' has taken place fairly recently and the knowledge about the use of saliva as a representative biological fluid has increased over the past years. Salivette samples are obtained by placing a small cotton swab into the mouth for two minutes after which it is stored in a closed plastic container. Because total serum cortisol may be altered by fluctuations in binding proteins, free serum cortisol is a better indication of adrenal activity. Salivary cortisol measurements always reflect the biological active free form. Salivary free cortisol is approximately 70% of that of serum free cortisol because of conversion of cortisol to cortisone in the salivary glands. However, there is a strong relationship between cortisol levels extracted from saliva and from blood (Riad-Fahmy *et al.*, 1982; Kirschbaum and Hellhammer 1994; Aardal and Holm 1995). In urine cortisol exists only in free form; secretion is dependent on serum levels, but also on renal glomular and tubular function. Both blood and saliva can provide information on the diurnal rhythm, while urine measures represent the cortisol production over a period of time. The latter is less informative, but, because it is a summary index, may show better reliability than blood samples. A possibility to get informative urine samples, taking the diurnal cycle into account, is by collecting at different times of the day instead of the commonly used 24-hour pooling method.

All three methods have their pro's and cons and it depends on the aim of the studies which method is more appropriate (Riad-Fahmy et al; 1982; Kirschbaum and Hellhammer, 1989; Trainer *et al.*, 1993; Aardal and Holm 1995; Kirschbaum and Hellhammer 1994). Likewise, there is no theoretically optimal measure suitable for twin studies, because there too it depends on what exact cortisol phenotype is of interest. In the existing studies, significant genetic influences on baseline as well as stimulated cortisol levels have been found in saliva (Kirschbaum *et al.*, 1992; Young *et al.*, 2000; Wüst *et al.*, 2000;), blood (Maxwell *et al.*, 1969; Meikle *et al.*, 1988; Linkowski *et al.* 1993; Inglis *et al.*, 1999) and urine (Inglis *et al.*, 1999). The difference in collection method may explain part of the difference in heritability estimates, but cannot explain all of it. For example, both the study by Froelich *et al.* (2000) and Inglis *et al.* (1999) used plasma samples, collected in the morning, and analyzed by making use of a RIA. The two studies, nonetheless, find severely discrepant results with Froelich *et al.* showing no sign of heritability and Inglis *et al.* reporting a heritability estimate of 46%.

Methods of analysis

Levels of cortisol in plasma, urine and saliva can be estimated by commercial radioimmunoassay (RIA), high performance liquid chromatography (HPLC), and enzyme-linked immunosorbent assay (ELISA). All three methods of analysis have their pro's and

cons and it again depends on the aim of the studies which method is more appropriate (Liddle, 1960; Kuhn, 1989; Okumura et al, 1995). The commercial availability of sensitive and specific antisera for cortisol has made RIA the method of choice in most laboratories. However, the advantage of using an ELISA is that it does not require the use of any radioactive reagent, and therefore it is safer and more economical than standard RIAs. The disadvantage is that the sensitivity of an ELISA does not quite approach that available with standard RIA. HPLC might be more specific and accurate, but they are also considerably more time consuming, and require much more complicated instrumentation. There is again no theoretical advantage of any method in twin studies.

However, what is important is the handling of batch effects. Perhaps counter-intuitively, a random distribution of the samples of family members over different batches is required to avoid confusing the genetic experiment. Like any family study, the classical twin design can provide evidence of familial aggregation of cortisol levels. In addition, the twin study exploits the difference in genetic similarity between two members of an MZ and of a DZ twin pair to differentiate between the two factors of familial aggregation: genetic influences and shared environmental influences. However, a third factor accounts for the observed variance in cortisol, the so-called nonshared environmental influences. These are influences unique to an individual, and include unpredictable measurement error due to the distance of the sample to the last CRF pulse but also predictable measurement error due to batch effects. If the within-family batch effects are removed, but between family batch effects are left to exist (because not all families can be run in a single batch), the extent of familial aggregation will be overestimated in MZ as well as DZ twin pairs. This overestimation will show up as influences of shared environmental factors. Random distribution over batches will force the batch effect to show up as a nonshared environmental effect, which is appropriate.

None of the studies reviewed in Table 9.1 mentioned a random distribution over different batches. Hence, it is possible that samples of two members of a twin pair are analyzed in the same batch as a result of sample collection on the same day. Random distribution of the samples is the procedure to avoid correlated measurement errors, which are difficult to detect but could certainly influence the results.

Timing of sampling

The secretion of cortisol is a prominent part of the endocrine response to stress. Because of the complexity of the HPA-axis mechanism and the discrepancy in physiological background of basal cortisol levels in comparison to cortisol release in response to a stressor (psychological stress, chemical substance, exercise), we will focus on basal cortisol levels solely. Six of the 11 studies (Nurnberger *et al.*, 1982; Kirschbaum, 1992;

Pritchard *et al.*, 1998; Froehlich *et al.*, 2000; Inglis *et al.*, 1999; Pritchard *et al.*, 1999) focus on the cortisol response to a certain stimulus, but also took basal samples to determine baseline cortisol levels before application of the stressor. Unfortunately, because their basal sample only acted as baseline for computation of the reactivity levels, most completely ignored the timing of the basal sample in the circadian rhythm. Four out of these 6 studies did not find evidence for genetic influences on the basal levels. Only Inglis *et al.* (1999) and Kirschbaum *et al.* (1992), report significant genetic influences on basal cortisol.

Five of the 11 studies (Maxwell *et al.*, 1969; Meikle *et al.*, 1988; Linkowski *et al.*, 1993; Young *et al.*, 2000; Wüst *et al.*, 2000), listed in Table 9.1, focus specifically on basal cortisol levels. These studies show an unfortunately large variation in frequency and the timing of the sampling across the measurement day. Some studies sampled twice or more a day, at fixed hours and over a longer period, whereas others sampled only once a day and not even at fixed times. Clearly, based on the knowledge of the circadian rhythm, the frequent sampling at fixed times is favored and except for those studies that used urinary sampling to measure 24 hour cortisol profiles, the other approaches introduce large between-subject variance due to time of sampling. 24-hour averages have the disadvantage, however, that they assume the cortisol level at all time points of the day to be influenced by the same genetic or environmental influences. This assumption need not hold, in the view of the complexity of the HPA system. It is entirely possible that different genes influence cortisol at different times of day.

In summary, the differences in the estimates for genetic influences could in part be due to collection methods, handling of batch effects, different time schedules for sample collection and different focus of studies (basal or reactivity). Apart from these differences in methodology, however, a major problem in most studies is the small sample size. The statistical power of quantitative genetic studies is influenced by the size of the effect (e.g. heritability), the sample size, the probability level (α) chosen, and the homogeneity of the sample (see among others, Neale and Cardon, 1992). Table 9.1 clearly shows that the number of twin pairs used in the different studies is rather low. This may be sufficient to demonstrate familial effects, but the statistical power to distinguish between genetic or shared environmental influences (environmental influences shared by different members of a twin pair or family) as the primary cause of familial aggregation may still be insufficient.

A SIMULTANEOUS ANALYSIS

To deal with this problem of small sample sizes, we performed a simultaneous analysis on those five studies that used more or less comparable methodology to measure basal

cortisol levels and which provided description of the sample size and MZ and DZ correlations. The studies used in the simultaneous analysis are described in Table 9.2.

Table 9.2.

Descriptives of the studies used in the simultaneous analysis

Study	**Sample**	**Sample Size**	**MZ and DZ Correlations**
1. Wüst *et al.*, 2000	saliva	52 MZ 52 DZ	rMZ: .58 rDZ: .38
2. Froehlich *et al.*, 2000	plasma	51 MZ 37 DZ	rMZ: .57 rDZ: .49
3. Inglis *et al.*, 1999	plasma	75 MZ 71 DZ	rMZ: .50 rDZ: .27
4. Linkowski *et al.*, 1993	plasma	11 MZ 10 DZ	rMZ: .59 rDZ: .60
5. Meikle *et al.*, 1988	plasma	20 MZ 20 DZ	rMZ: .50 rDZ: .24

What is immediately evident is that the MZ correlations of all studies are approximately .55, whereas the DZ correlations vary from .24 to .60. Based on the stable MZ correlation it was expected that basal cortisol levels are influenced by genetic factors with a maximum heritability of 60%. In absence of shared environmental influences the MZ correlation is equal to the heritability. However, MZ correlation alone cannot distinguish between genetic or shared environmental influences as the primary cause of familial aggregation. Making use of the difference in MZ and DZ correlation, a simultaneous analysis was used to disentangle these sources. Structural equation modeling (Mx, see Neale *et al.*, 1999) was used to fit, by maximum likelihood estimation, the observed twin correlations of all studies against different theoretical models (Table 9.3). The full model allows for genetic (A), shared environmental (C), as well as nonshared environmental (E) influences on cortisol. More parsimonious models then leave out the genetic or environmental influences and test the loss of fit to the observed data by calculating the change in χ^2 ($\Delta\chi^2$) against the gain of degrees of freedom (Δdf). First, it was tested whether the five studies could be taken together to use all information to estimate the genetic and environmental influences (model 2, 3, 4). Taking the various studies together did not result in a significant change of fit. Secondly, the significance of shared environmental influences was tested by dropping this factor from the model

(model 5 and 6), which did not result in a significant change of χ^2. However, dropping genetic influences did significantly change the fit of the model (model 7 and 8).

Table 9.3.

Model-fitting results of the simultaneous analysis based on twin correlations from 5 studies (see table 9.2).

Model	χ^2	df	χ^2	df	c.t.m.[g]	p	A	C
1. AC[a], all different[b]	.009	0					.00-.49[c]	.00-.59
2. AC, A equal	1.05	4	1.04	4	1	n.s[d]	.40	.05-.32
3. AC, C equal	3.16	4	3.07	4	1	n.s	.31-.65	.13
4. AC, A equal, C equal	7.89	8	7.88	8	1	n.s.	.37	.17
5. A different, no C	5.36	5	5.35	5	1	n.s.	.49-.88	-
6. A equal, no C[e]	11.54	9	6.18	4	5	n.s.	.62	-
7. C equal, no A[f]	16.30	9	13.14	5	3	.02	-	.45
8. C different, no A	11.69	5	11.68	5	1	.04	-	.37-.59

[a] A= additive genetic factors, C= common (shared) environmental factors; [b] different estimates for the separate studies; [c] range of estimates for the different studies; [d] n.s. = non-significant determination of fit; [e] Common environmental influences are dropped from the model; [f] additive genetic influences are dropped from the model; [g] c.t.m. = compared to model

POWER ANALYSIS

Based on the simultaneous analyses it can now be concluded that genetic factors are the major source of the familial aggregation. No evidence was found for shared environmental factors, but a major question remains whether statistical power, even with pooling of studies, was sufficient. A third and final purpose of this study, therefore, was to estimate the number of twin pairs required to obtain reliable estimates of heritability and shared environmental variance. A power analysis using Mx (Neale *et al.*, 1999) was conducted to calculate the required sample sizes given heritability estimates of 5%, 15%, 25% 35%, 45% and 55%, so that rMZ=.60 (Accordingly estimates of shared environmental influences are 55%, 45%, 35%, 25%, 15% and 5%, respectively). The assumption of rMZ=.60 was made based on the MZ correlations used for the simultaneous analyses.

The analyses were again based on a comparison of different models. To estimate the sample size to detect genetic influences a model with additive genetic (A), shared environmental (C) and nonshared environmental (E) factors was compared to a model with shared environmental and nonshared environmental factors only. Further, to estimate the sample size needed to detect the shared environmental influences, an ACE model was compared to an AE model. Finally, to estimate the sample size for detection

of familial aggregation an ACE model was compared to a model with nonshared environmental influences only. Table 9.4 shows the required number of twin pairs on the basis of the power desired. As can be seen in Table 9.4, small effect, i.e. low heritabilities, are harder to detect and need far larger sample sizes than have been used in the existing studies.

Table 9.4.

Sample sizes needed to estimate genetic influences, common environmental influences and familial aggregation with a power of .80 and with simulated twin data (equal number of MZ and DZ).

Variance distribution	**detection of genetic influences**	**detection of common environmental influences**	**detection of familial aggregation**
h^2=.05, c^2=.55[a]	16087[b]	136	19
h^2=.15, c^2=.45	1984	215	21
h^2=.25, c^2=.35	786	374	22
h^2=.35, c^2=.25	437	769	24
h^2=.45, c^2=.15	286	2231	26
h^2=.55, c^2=.05	206	20878	28

[a] h^2 = heritability, c^2 = common environmental influences; [b] total number of complete twin pairs (equally divided in MZ and DZ).

DISCUSSION

The first purpose of this chapter was to critically examine the existing literature on the heritability of cortisol levels in twin and family studies. We found 11 studies that satisfied our search criteria; based on the search terms *twin, cortisol, corticosteroid, heritability* and *family*, no case studies, published in English, and genetically related subjects (see introduction). After careful inspection we concluded that these studies lack the methodological consistency required for a good comparison. A factor that makes it hard to compare the different studies is the multitude of cortisol measures used: basal cortisol, area under the curve, morning peak, nocturnal nadir and reactivity to acute stressors. These are clearly different phenotypes and need not be influenced by the same genes or environmental factors. The studies that focus on basal levels have used rather equivocal frequency and timing of sampling across the day. Also where the older studies sometimes neglected confounding factors such as physical exercise, smoking or type of depression, the more recent studies often rule out all possible individual differences by selecting non-smoking, same-sex subjects that use no oral contraceptives or other over the counter medications.

This is good experimental practice, but makes it hard to compare them to the older studies. It is also uncertain how this selection affects the generalisability of the genetic architecture to that of the general population. Twin pairs with cortisol levels outside the expected range are mostly excluded, for instance, while it might be interesting to know why those twin pairs deviate.

The main problem plaguing existing studies is the relatively low number of twin pairs on whom cortisol was obtained. A power analysis revealed that none of the 11 studies examined consists of a large enough sample size to be able to separate genetic and environmental influences. A combined analysis of these studies estimated the heritability of 62% for basal cortisol levels with a combined sample size of 399 twin pairs (209 MZ and 190 DZ). According to Table 9.4, this simultaneous analysis has the power to separate familial influences into shared environmental and genetic influences. Disentangling the sources of familial aggregation is essential in understanding individual differences in basal cortisol levels. Previous studies point out that both genetic factors (see Table 9.1) and shared environmental factors, such as parenting styles, influence basal cortisol levels.

The obvious approach to increase power is to increase the sample size. The recent development of large Twin Registries all over the world (e.g. Boomsma, 1998) and of ambulatory sample collection methods will make it more feasible to measure on large numbers of subjects relatively easily in the future. Even so, the sample sizes shown in Table 9.4 are in some cases unrealistic high and beyond the scope of actual resources of time, money and practical attainability. Other remedies, therefore, must be considered. One might be the composition of the sample. Increase of power can be achieved by extending the twin design by adding siblings (Posthuma and Boomsma, 2000) or adding information from adoption studies (Schmitz *et al.*, 1998). Most importantly, an increase in power can be achieved through multivariate analyses, i.e. by increasing the number of measurements through repeated measure designs. Provided that those repeated measurements correlate significantly with each other, this yields large gains in power (Schmitz *et al.*, 1998).

Increasing the number of samples on a single measurement day is a first method to obtain repeated measures. However, in the view of the complexity of the HPA system, it is entirely possible that different genes influence cortisol at different times of day. It is valuable as such to assess the genetic architecture of cortisol level at different points of the diurnal curve, i.e. to understand the sources underlying individual variation in the morning peak level as well as in the evening. But if the samples across different time points are not (genetically) correlated increasing the number of samples on the same measurement day will do little to improve power to detect heritability. The optimal

approach is to sample at *repeated time points* of the diurnal curve, on *repeated days* within the same subject as was done by Young *et al.* (2000). This still leaves error variance due to differences in day-to-day variation in wake-up times, but this can be dealt with by sampling at fixed time points relative to wake-up times.

In summary, understanding the genetic architecture of basal cortisol level awaits studies with large twin samples that measure cortisol repeatedly at fixed time points from the awakening time, and do so on repeated days. In parallel, twin studies must be aware of the rapid progress in animal research. A number of candidate genes with respect to basal cortisol levels have emerged, like those that affect corticotrophin-releasing hormone (CRH) or ACTH synthesis (and also the production of their receptors) and those that code for mineralocorticoid (MR, or type-I receptor) and the glucocorticoid (GR, or type-II receptor). Polymorphism(s) in the latter gene have already been associated with various aspects of cortisol metabolism such as varying basal cortisol levels (Rosmond *et al.*, 2000a; Rosmond *et al.*, 2000b) and differences in sensitivity to glucocorticoids (Huizinga *et al.*, 1998). One of the huge advantages of a twin sample is that such observed genes (candidate genes) and unobserved genes (estimates of heritability through the MZ-DZ comparison) can be simultaneously tested.

10

Heritability of Daytime Cortisol Levels in Children

This chapter is under revision as:

Bartels, M., de Geus, E. J.C., Sluyter, F., Kirschbaum, C., and Boomsma, D.I. (2002). Heritability of Daytime Cortisol in Children. *Behavior Genetics*, under revision

ABSTRACT

Individual differences in the level of the stress hormone cortisol levels play a prominent role as an explanatory variable in studies on psychopathology. Relatively few studies have paid attention to individual differences in cortisol levels and the etiology of these differences, in particular their possible genetic basis. All these studies have been in adults. The aim of this study was to estimate genetic and environmental influences on basal cortisol levels in twelve-year-old children. To this end, four samples of salivary cortisol were collected on two consecutive days in a sample of 180 twin pairs. Low correlations were found between cortisol levels at different points in time during the day. A significant genetic contribution was found to the variation of basal cortisol levels in the morning and afternoon samples, but not in the evening sample. Heritability did not differ for boys and girls and was highest (60%) for cortisol levels during the sample taken about 45 minutes after awakening. This cortisol awakening response provides a useful endophenotype in the search for genes that may affect hypothalamic-pituitary adrenocortical functioning in children.

INTRODUCTION

Cortisol is a steroid hormone secreted by the outer cortex of the adrenal gland. Its secretion is stimulated by ACTH (adrenocorticotrophic hormone), produced in the pituitary in response to corticotropin-releasing hormone (CRH), a product from neurons in the paraventricular nucleus of the hypothalamus. After its release, the major part of cortisol binds to the plasma proteins corticosteroid binding globulin (CBG or transcortin) and albumin, which prevents the hormone from penetrating the membranes of their target cells. About 3-5% of the total cortisol is the unbound, biologically active fraction. This active fraction has permissive, suppressive, stimulatory, and preparative action effects in the realms of cardiovascular function, fluid volume and hemorrhage, immunity and inflammation, metabolism, neurobiology, and reproductive physiology (Sapolsky, Romero & Munck, 2000). Although cortisol is mainly known for its pivotal role in generating an adequate response to physical and emotional stressors, it may also exert strong behavioral effects that are already apparent during childhood. Many studies have reported an association between cortisol levels and Internalizing and Externalizing problem behaviors in children (McBurnett *et al.*, 2000; 1996; 1991; Dawes *et al.*, 1999; Van Goozen *et al.*, 1998; Scerbo and Kolko, 1994; Vanyukov *et al.*, 1993; Tennes and Krey, 1985; Dorn *et al.,* 1999; Granger *et al.*, 1994; Scerbo and Kolko, 1994; McBurnett *et al.*, 1991; Kagan *et al.*, 1987; Tennes *et al.*, 1986). An obvious question for behavior geneticists, therefore, is whether the known genetic contribution to these problem behaviors is partly mediated through genetic effects on the hypothalamic-pituitary adrenocortical (HPAC) axis generating this important stress hormone.

In the characteristic diurnal rhythm of plasma cortisol level, typically 10-15 well-defined pulses of variable amplitude are observed, with a morning maximum, declining levels throughout the daytime, a period of low concentrations generally centered around midnight, and an abrupt rise after the first few hours of sleep (Weitzman, 1971). Within the first 30 minutes after awakening, free cortisol levels rise by 50-60% (Pruessner *et al.*, 1997; Wüst *et al.*, 2000). Plasma cortisol release is tightly regulated through negative feedback at the pituitary, hypothalamus and hippocampus (Kovacs *et al.*, 1987; Jacobson & Sapolsky, 1991). This negative feedback is mediated via two types of adrenal steroid receptors: the high-affinity mineralocorticoid receptors (MR) in the hippocampus and the low-affinity glucocorticoid receptors (GR) widely distributed throughout the brain. Strength of this feedback signal strongly varies with time of day (Dorin *et al.*, 1996; Huizinga *et al.*, 1998; Young *et al.*, 1998), contributing to the characteristic diurnal rhythm in plasma cortisol levels (See Figure 10.2). Because the activated GR and MR receptors act as *trans*acting factors (Meyer, de Kloet and McEwen, 2000) it is likely that genetic variation in the *cis*acting elements for these activated receptors can act to create significant

individual variation in diurnal cortisol profiles. However, genetic variation in cortisol levels may also arise at many other points in HPAC axis, for instance in the synthesis of corticotrophin-releasing hormone (CRH) or ACTH or in the production of their receptors or those that code for mineralocorticoid (MR, or type-I receptor) and the glucocorticoid (GR, or type-II receptor) receptors themselves. In animal studies, polymorphism(s) in the latter gene have already been associated with various aspects of cortisol metabolism such as varying basal cortisol levels (Rosmond *et al.*, 2000a; Rosmond *et al.*, 2000b) and differences in sensitivity to glucocorticoids (Huizinga *et al.*, 1998).

Twin studies constitute a powerful method for identifying genetic influences on (diurnal changes in) cortisol levels in humans. Surprisingly few attempts have been made to estimate the relative impact of genetic and environmental factors on the regulation of cortisol levels (for a review see Bartels *et al.*, 2002c). Most of these studies point to the direction of moderate genetic contributions to different aspects of cortisol measures.

Thus, Maxwell and colleagues (1969) showed a significant smaller intrapair variance in MZ as compared to DZ female twin pairs. Meikle *et al.* (1988) reported evidence of moderate genetic effects (h^2= 51%) on basal cortisol levels in males. More recently, Inglis *et al.* (1999) has reported a heritability of 46% in morning plasma cortisol samples. Furthermore, significant heritabilities have been found in the cortisol stress-response and in the cortisol response to awakening (Kirschbaum *et al.*, 1992; Wüst *et al.*, 2000). The main problem plaguing many of these studies is the relatively low number of twin pairs on whom cortisol was obtained. A power analysis revealed that none of the previous studies on the heritability of cortisol levels consists of a large enough sample size to be able to distinguish between genetic or shared environmental influences (environmental influences shared by different members of twin pair or family) as the primary cause of familial aggregation (Bartels *et al.*, 2002c).

The obvious approach to increase power is to increase the sample size. The recent development of large Twin Registries all over the world (e.g. Boomsma, 1998) and of ambulatory salivary sample collection methods (Aardal and Holm, 1995; Kirschbaum and Hellhammer, 1994; Riad-Fahmy *et al.*, 1982) has made it more feasible to measure on large numbers of subjects relatively easily in the future. Actual resources of time, money and practical attainability now mostly restrict the sample size. The power of twins studies to detect additive genetic or environmental variation, however, can be increased through other means besides increasing the sample size (Cohen, 1977; Neale *et al.*, 1994). Relevant for the present study is that an increase in power can be achieved through multivariate analyses, for instance by a repeated measurement design. Provided that those repeated measurements correlate significantly with each other, this yields large gains in power (Schmitz *et al.*, 1998). In this multivariate method not only the expectation for the within-

pair covariances is taken into account but also the cross-trait as well as the within-person information. A first method to obtain repeated measures on basal cortisol is to repeatedly sample across an entire measurement day.

A first practical strategy would be to sample repeatedly on a single measurement day. However, there are large changes in mean and variance due to the circadian rhythm. It is entirely possible that the contribution of genetic variance changes across the day, possibly due to the expression of different genes. From a physiological "content" point of view it is valuable per se to assess the genetic architecture of cortisol level at different points of the diurnal curve, i.e. to understand the sources underlying individual variation in the morning peak level as well as in the evening through. The optimal approach, therefore, is to sample the same time point across multiple days instead of multiple time points on a single measurement day.

Full understanding of the genetic architecture of basal cortisol level awaits studies with large twin samples that measure cortisol repeatedly at fixed time points from the awakening time, and do so on repeated days. Moreover, all our current knowledge on the genetics of cortisol comes from studies in adults. Estimates of the strength of genetic and environmental influences on variation in basal cortisol levels obtained in adults cannot be generalized to children. The developmental trajectories of the various steroid hormones are intertwined and points of cross-talk between the HPAC axis and the gonadal hormones have been shown (Vamvakopoulos and Chrousos, 1994). Just by considering the large changes in gonadal hormone levels from childhood to adolescence, it would be unwise to extrapolate adult genetic architecture of cortisol levels (which itself is unlikely to be stable across the adult life span) to pre-adolescent children. Indeed, for the behavioral phenotypes possibly influenced by cortisol, cognitive ability and problem behavior, a change in the strength of genetic and environmental influences throughout development has already been observed (Bartels *et al.*, 2002b, 2002e). So, insight into the cause of individual differences in basal cortisol levels in childhood, besides the current knowledge in adults, is essential.

The aim of this study is to determine the heritability of variation in daytime cortisol levels in children. In accordance with methodological consideration mentioned above, we collected saliva samples on four fixed points of time on two consecutive days in a large group of twelve-year-old twins.

METHODS

Subjects

This project is part of an ongoing, longitudinal study on the development of cognition and emotional and behavioral problems in children. Details on the demographic characteristics

of the sample and information on parental occupation can be found elsewhere (Rietveld *et al.*, 2000). For the determination of cortisol levels, saliva was collected in 1999/2000 when the twins were 12 years old. Mean age of the subjects was 12 years (80% ranging from 11 years and 11 months to 12 years and 1 month). Zygosity of the same-sex twins was established by either blood group polymorphisms (137 pairs) or DNA analyses (24 pairs) and in nine pairs by physical resemblance assessed by an experienced test-administrator. The twin sample at age 12 consisted of 47 monozygotic female (MZF), 37 dizygotic female (DZF), 42 monozygotic male (MZM), 44 dizygotic male (DZM), and 39 dizygotic pairs of opposite sex (DOS). Because of difficulties during saliva collection or laboratory analyses, the final sample consisted of 180 twin pairs. The exact numbers of cortisol samples for each point in time can be found in Table 10.1. Pubertal status has been determined by the Tanner scales.

Saliva collection

Four samples of cortisol per day on two consecutive days were collected using the Salivette sampling device (Starstedt, Rommelsdorf, Germany). Salivary cortisol measurements reflect the biological active free form. Salivary free cortisol is approximately 70% of that of serum free cortisol because of conversion of cortisol to cortisone in the salivary glands. However, salivary cortisol levels correlate very strongly with plasma free cortisol (Aardal & Holm, 1995, Kirschbaum & Hellhammer, 1994, Riad-Fahmy *et al.*, 1982).

Salivettes were sent to the participants by mail and the twin pairs collected their saliva at home, following a written instruction. The samples were collected at prescribed times and, importantly, at the same time for both children of a twin pair. On the first day the first sample (S11) was taken in the morning just before getting up (still lying in bed) (mean time 0728H), the second (S21) sample was taken at least half an hour after getting up but before going to school (mean time 0817H), the third sample (S31) was taken before lunch (mean time 1234H), and the fourth sample (S41) was taken in the evening (mean time 2032H). On the second day the same schedule was adapted for four repeated samples (S12, S22, S32, S42)(mean times 0735H, 0826H, 1232H, 2034H). The twins were instructed to collect saliva on two school days to restrict the awakening time and time of sampling. School starting time and lunch break is at approximately the same time all over the Netherlands resulting in small sampling time variation. Each participant was asked to write down the exact sampling time in a time-schedule and to note exceptional events interfering with daily routine. Subjects were instructed not to brush their teeth before completing saliva sampling to avoid contamination of saliva with blood caused by micro-injuries in the oral cavity. Also, subjects were instructed to thoroughly rinse their mouth

with tap water before sampling saliva and not to eat sour food or drink aerated drinks. Subjects were strictly instructed to collect saliva before taking lunch at time point three (S3). Saliva samples were stored in the freezer until completing the experimental protocol and the samples were picked up by the test-administrator and sent by courier to the laboratory in Germany (Trier and Düsseldorf).

Saliva sampling

The saliva samples of twins of the same pair were randomly distributed over different batches, but the samples of a single subject were placed in one batch. The analyses were performed without knowledge of the zygosity of the twins and without knowledge of exact time of collection. Saliva samples were spun at 3300 rpm for five minutes, and cortisol in saliva was determined by time-resolved fluorescence immunoassay, as described elsewhere (Dressendörfer *et al.*, 1992; Wüst *et al.*, 2000). Intra- and interassay variability of the assay was less than 10 and 12%, respectively.

Data analyses

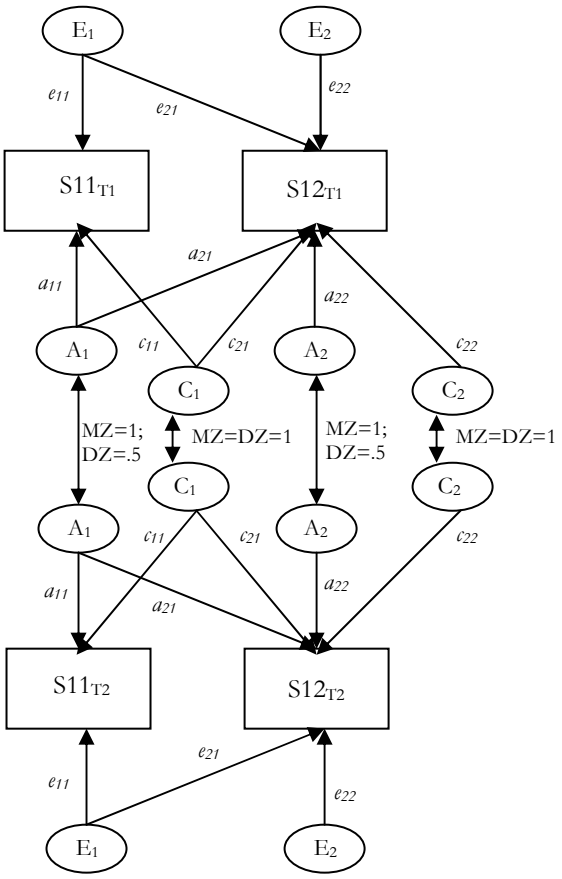
Descriptive statistics for S11 (sample one at day one), S21, S31, S41, S12, S22, S32, and S42 were calculated using SPSS/windows 10. Pearson correlations were used to test the association between the samples collected on the same day and the association between the samples taken at the same point in time on the two consecutive days. MZ and DZ cross correlations and twin correlations for the five zygosity groups (MZM, DZM, MZF, DZF, DOS) have been calculated to get a first impression of the genetic and environmental influences on salivary cortisol levels at the different points in time. The cross-correlations represent cross-day-cross-twin correlation and in that matter represents the repeated measurement design. For instance, sample one, at day one for the oldest of the twin is correlated with sample 1, at day 2 for youngest of the twin.

Genetic Modeling

Genetic model fitting of twin data allows for separation of the observed phenotypic variance into its genetic and environmental components. Additive genetic variance (A), is the variance that results from the additive effects of alleles at each contributing genetic locus. Shared environmental variance (C) is the variance that results from environmental events common to both members of a twin pair. Unique environmental variance (E) is the variance that results from environmental effects that are not shared by members of a twin pair. Estimates of the unique environmental effects also include measurement error. To account for this source of variance, E is always specified in the model.

Figure 10.1.

Cholesky Decomposition Model for a sample on the first day and a sample on the second day at the same time



The different degree of genetic relatedness between monozygotic (MZ) and dizygotic (DZ) twin pairs was used to estimate the contribution of these factors to the phenotypic variation in cortisol levels (Plomin *et al.*, 1997). Similarities for MZ twins are assumed to be due to additive genetic influences plus environmental influences that are shared by both members of a twin pair. Experiences that make MZ twins different from one another are unique environmental influences. Because DZ twins share 50% of their genetic material on average, like other siblings, genetic factors contribute only half to their resemblance. As for MZ twins the shared environment contributes fully. Model fitting to twin data is based on the comparison of the variance-covariance matrices in MZ and DZ twins. Exploiting the known difference in genetic contribution to intra-pair resemblance of MZ and DZ

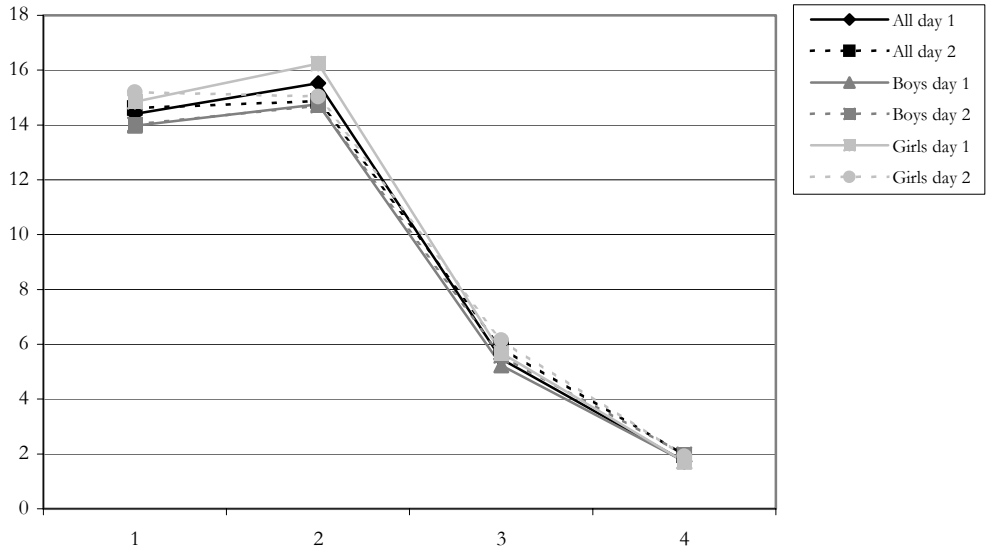
twin pairs, influences of additive genetic, shared environmental and unique environmental factors are estimated using the computer program Mx (Neale *et al.*, 1999).

Per time point a bivariate model (Cholesky decomposition), based on cortisol samples from the same point in time on the two consecutive days, was used to estimate genetic and environmental influences (see Figure 10.1). Rather than decomposing the variance of a single cortisol sample into genetic and environmental sources of variance, bivariate genetic analysis decomposes the variance of each sample and the covariance between the samples at the same time on the two measurement days into genetic and environmental sources.

To make optimal use of all available data, including incomplete twin pairs, analyses were performed on the raw data. In Mx the handling of such 'incomplete' data is implemented by calculating twice the negative log-likelihood (-LL) of the raw data of each twin pair and sum these over all pairs. When two models, which provide –2LLs, are nested subtracting the two –2LLs from each other provides a (-2LL), which has a χ^2 distribution. A high χ^2 against a low gain of degrees of freedom (df) denotes a worse fit of the second, more restrictive model relative to the first model. If no significant difference is observed, the more parsimonious model is preferred.

Figure 10.2.

Graphical representation of the mean cortisol levels (nmol/L) at each measurement occasion for the total sample and boys and girls separately.



We began with fitting an ACE model. The issue of possible sex-differences in heritability or environmental influences is sorted out in the model fitting procedures. First, we tested whether different genes influence basal cortisol levels in boys and girls or whether the same or different shared environmental factors influence cortisol levels in boys and girls. It was also tested whether the influences of the genes are of different strength in boys and girls. Significance of genetic and shared environmental influences was tested. To this end it was tested whether a model with additive genetic and unique environmental influences only (AE), gave a significantly worse fit than the full model (ACE). It was also tested whether a model with shared environmental influences and unique environmental influences (CE) gave a significantly worse fit than the full model (ACE). Finally, it was tested whether individual differences of cortisol levels are based on unique environmental influences, solely (E model).

Estimates of genetic, shared environmental and unique environmental influences on basal cortisol levels at each point of time separately have been estimated based on the best fitting model. Because the between time point (on the same day) correlations were very low, multivariate models with different cortisol samples taken on the same day but at a different point in time were not considered meaningful.

Results

Descriptive statistics of the cortisol measures assessed during the day are presented in Table 10.1. Means and standard deviations have been calculated for the entire sample and for boys and girls separately. Skewness and Kurtosis showed that the variables were approximately normal distributed, so no transformation was conducted. Figure 10.2 shows the expected circadian rhythm with an increase of cortisol levels in the morning and decreasing levels over the day. No significant differences for boys and girls are observed. No significant influence of pubertal status on cortisol levels could be observed.

Phenotypic correlations are presented in Table 10.2. Very low associations are found between samples taken on the same day at different time points. Significant correlations are found between samples taken at the same point in time on the two consecutive days (boldfaced). The MZ and DZ cross-correlations, presented in Table 10.3, suggest influences of genetic factors on the association between the two same samples on the two different days (boldfaced). Twin correlations for the five zygosity groups separately are presented in Table 10.4. As for the MZ and DZ correlation, these twin correlations suggest genetic influences on S1, S2, and S3. Individual variation for S4 is mainly due to environmental factors.

Table 10.1.

Descriptives of cortisol (nmol/L) for all subjects together and for males and females separately.

		N[a]	Min	Max	Mean	Std	Skewness		Kurtosis	
								s.e.		s.e.
S11[b]	**all**	309	5.42	27.86	14.40	4.66	.566	.139	-.211	.276
	♂	158	5.78	27.35	13.97	4.66	.675	.193	.165	.384
	♀	151	5.42	27.86	14.85	4.62	.476	.197	-.216	.392
S21	**all**	324	2.38	10.19	5.46	6.69	.672	.135	.174	.270
	♂	158	2.43	10.19	5.23	6.15	.615	.193	.165	.384
	♀	166	2.38	9.50	5.69	7.10	.646	.188	.015	.375
S31	**all**	315	2.38	10.19	5.46	1.66	.429	.137	-.465	.274
	♂	155	2.43	10.19	5.23	1.67	.674	.195	.049	.387
	♀	160	2.38	9.50	5.69	1.63	.224	.192	-.749	.381
S41	**all**	293	.59	3.98	1.73	.67	.892	.142	.395	.284
	♂	150	.59	3.78	1.73	.67	.892	.198	.425	.394
	♀	143	.62	3.98	1.72	.66	.901	.203	.418	.403
S12	**all**	317	6.07	24.96	14.62	4.53	.323	.137	-.759	.273
	♂	158	6.07	24.96	14.02	4.31	.549	.193	-.317	.384
	♀	159	6.13	24.96	15.20	4.68	.106	.192	-.969	.383
S22	**all**	293	5.14	27.99	14.87	5.44	.388	.142	-.671	.284
	♂	146	5.14	26.51	14.71	5.25	.417	.201	-.628	.399
	♀	147	5.23	27.99	15.03	5.65	.356	.200	-.714	.397
S32	**all**	309	2.07	11.27	5.87	1.93	.445	.139	-.337	.276
	♂	149	2.30	10.64	5.58	1.81	.499	.199	-.325	.395
	♀	160	2.07	11.27	6.15	2.01	.355	.192	-.386	.381
S42	**all**	296	.53	4.87	1.95	.89	.961	.142	.619	.282
	♂	142	.55	4.79	1.99	.86	.912	.203	.762	.404
	♀	154	.53	4.87	1.91	.92	1.028	.195	.603	.389

[a] the total number of children ; [b] the first number refers to the time of sampling and the second number refers to the day; S11= first sample at day one; the cross correlation between the same time of sampling at the two different days is bold faced.

Table 10.2.

Phenotypic cross-correlations of cortisol levels with their 95% confidence interval

	S11[a]	S21	S31	S41	S12	S22	S32
S21	.02 *(-.10-.15)*	-					
S31	.21 *(.10-.33)*	.22 *(.10-.33)*	-				
S41	.03 ***(-.10-.15)***	.03 *(-.09-.15)*	.09 *(-.03-.21)*	-			
S12	**.36 *(.25-.46)***	.10 *(-.02-.21)*	.14 *(.03-.25)*	.10 *(-.03-.23)*	-		
S22	.01 *(-.12-.14)*	**.36 *(.24-.46)***	.23 *(.10-.33)*	.02 *(-.11-.14)*	.03 *(-.08-.15)*	-	
S32	.07 *(-.04-.19)*	.19 *(.07-.30)*	**.24 *(.13-.35)***	.14 *(.02-.26)*	.17 *(.05-.28)*	.21 *(.09-.32*	-
S42	.00 *(-.12-.12)*	.06 *(-.07-.19)*	.16 *(.04-.27)*	**.21 *(.09-.33)***	.10 *(-.02-.22)*	.18 *(.05-.29)*	.04 *(-.08-.16)*

[a] the first number refers to the time of sampling and the second number refers to the day; S11= first sample at day one; the cross correlation between the same time of sampling at the two different days is bold faced.

The results of the bivariate model fitting procedure for each time point demonstrate different contributions of genetic and environmental influences at the four cortisol measures (Table 10.5). However, no significant sex-differences have been found. For sample 1 (S1) no clear distinction could be made between genetic or shared environmental influences as the primary cause of familial aggregation, both model 5 (AE) and 6 (CE) are not significant different from a model with both A and C present (model 4). Reducing the additive genetic influences to one common influence on day one and day two did not significantly worsened the fit (model 7).The best-fitting model for samples S2 is a Cholesky decomposition model with additive genetic influences and unique environmental influences (model 5). For sample S3 the same pattern as for S1 was found. The best fitting model is a model with additive genetic and unique environmental influences (model 7). For sample 4 (S4) no factors of familial aggregation could be detected. The best fitting model is a model with unique environmental influences only (model 7).

Unstandardized and standardized estimates of genetic and environmental influences based on the best fitting models are presented in Table 10.6. Significant genetic influences are found for sample 1 (22%, 24%), sample 2 (56%, 59%) and sample 3 (30%, 21%). The heritabilities on the two consecutive days show slight differences. This is due to differences in unique environmental influences that change across days because they also contain day-specific measurement error. Since the total variance equals 100%, the differences in unique environmental influences are reflected in the small differences in heritability.

Table 10.3.

MZ (above diagonal) and DZ (below diagonal) correlation and cross correlations

	S11a	S21a	S31a	S41a	S12a	S22a	S32a	S42a	S11b	S21b	S31b	S41b	S12b	S22b	S32b	S42b
S11a	-	.01	.24	.01	.28	-.02	.10	.06	**.33**[a]	.00	.09	.07	**.18**[c]	.00	.00	.03
S21a	.02	-	.30	-.11	.00	.47	.23	.02	.00	**.64**[a]	.16	-.12	-.04	**.46**[c]	.17	.10
S31a	.21	.16	-	.03	.10	.40	.29	.35	.09	.16	**.45**[a]	.00	.23	.35	**.24**[c]	.19
S41a	.03	.17	.14	-	.11	-.18	.15	.18	.07	-.11	-.06	**.12**[a]	.14	-.18	-.11	**.00**[c]
S21a	.41	.21	.17	.11	-	-.03	.20	.11	.18	-.04	.23	.14	**.43**[a]	.03	-.05	.13
S22a	.04	.33	.10	.17	.10	-	.30	.27	.00	.46	.35	-.18	.03	**.62**[a]	.25	.10
S32a	.06	.19	.24	.17	.16	.15	-	.12	.00	.17	.24	-.11	-.05	.25	**.31**[a]	.04
S42a	-.06	.15	.03	.27	.09	.16	-.03	-	.03	.10	.19	.00	.13	.10	.04	**.14**[a]
S11b	**.17**[b]	-.08	.13	.01	.08	.06	-.07	.07	-	.00	.24	.01	.28	-.02	.10	.06
S21b	-.08	**.32**[b]	.02	.08	-.02	.20	.08	.03	.02	-	.30	-.01	.00	.47	.23	.02
S31b	.13	.02	**.25**[b]	.04	.18	.03	.02	.09	.21	.16	-	.03	.10	.40	.29	.35
S41b	.01	.08	.04	**.14**[b]	.00	.11	.04	.08	.03	.17	.14	-	.11	-.18	.15	.18
S12b	**.08**[d]	-.02	.18	.01	**.19**[b]	.00	.05	.03	.41	.21	.17	.11	-	-.03	.20	.11
S22b	.06	**.20**[d]	.03	.12	.00	**.36**[b]	.05	-.03	.04	.33	.10	.17	.10	-	.30	.27
S32b	-.07	.08	**.02**[d]	.04	.05	.06	**.15**[b]	-.01	.06	.19	.24	.17	.16	.15	-	.12
S42b	.07	.03	.09	**.08**[d]	.03	-.03	-.11	**.23**[b]	-.06	.15	.03	.27	.09	.16	-.03	-

[a] Twin correlations for monozygotic twins; [b] Twin correlations for dizygotic twins; [c] Cross correlations for monozygotic twins; [d] Cross correlations for dizygotic twins

Discussion

The purpose of this study was to estimate the genetic and environmental influences on the variation in basal cortisol levels at four different time points on two consecutive days in a large sample of 12-year-old children. Although previous studies have all used adult twins, the findings were very much in line with previous findings on the genetic architecture of urinary or salivary cortisol levels (Bartels *et al.*, 2002c). A significant genetic contribution to basal cortisol levels was found at three of the four time points sampled. Heritability did not differ for boys and girls and was highest (60%) for cortisol levels during the second sample taken about 45 minutes after awakening. A major contribution of unique environmental factors was found that dominated interindividual variation at all time points, save the second sample.

Table 10.4.

Twin correlations in cortisol level with their 95% confidence intervals.

	MZM [a]	**DZM**	**MZF**	**DZF**	**DOS**
S11 [b]	.28 *(.00-.58)*	.11 *(.00-.42)*	.42 *(.07-.67)*	.00 *(.00-.33)*	.11 *(.00-.46)*
S21	.45 *(.06-.71)*	.34 *(.01-.60)*	.67 *(.44-.82)*	.23 *(.00-.55)*	.14 *(.00-.50)*
S31	.54 *(.19-.75)*	.40 *(.03-.67)*	.43 *(.08-.67)*	.50 *(.13-.73)*	.00 *(.00-.31)*
S41	.00 *(.00-.30)*	.53 *(.10-.76)*	.21 *(.00-.62)*	.13 *(.00-.52)*	.04 *(.00-.44)*
S12	.17 *(.00-.51)*	.00 *(.00-.34)*	.62 *(.36-.79)*	.35 *(.00-.64)*	.18 *(.00-.50)*
S22	.53 *(.17-.76)*	.23 *(.00-.52)*	.68 *(40-.83)*	.37 *(.00-.67)*	.35 *(.00-.62)*
S32	.03 *(.00-.43)*	.32 *(.00-.61)*	.31 *(.00-.59)*	.10 *(.00-.49)*	.00 *(.00-.32)*
S42	.00 *(.00-.49)*	.40 *(.04-.66)*	.17 *(.00-.51)*	.19 *(.00-.54)*	.00 *(.00-.27)*

[a] MZM= monozygotic males, DZM= dizygotic males, MZF= monozygotic females, DZM= dizygotic females, DOS= dizygotic opposite sex;[b]the first number refers to the time of sampling and the second number refers to the day; S11= first sample at day one

Wüst and colleagues (2000) found a similar pattern of heritabilities, with a moderate to high heritability estimate for the cortisol response to awakening (40%) and low to non-significant heritability estimates for cortisol samples later that day, where unique environmental influences dominated. As suggested by Wüst *et al.* (2000), sleep is a period of very low differentiation in environmental influences that only kick in fully after awakening, and accumulate during the day, giving rise to a gradual increase in environmental variance. This could lead to a shift from genetic to environmental control over individual variation in cortisol levels. We explicitly tested in what way the changes in genetic architecture across time points reflected a change in the ratio of genetic and environmental variance. In contrast to the suggestion by Wüst *et al.*, we found that the relative increase in genetic variance at the second sample compared to the other samples was much more pronounced than the increase in environmental and total variance. Alternatively, therefore, we hypothesize that the heritability of cortisol levels varies inversely with the strength of the negative feedback signal exerted by cortisol at the GR and MR receptors. Changes in the strength of this feedback signal are reflected in changes in the absolute cortisol level, although time lagged, because the effects of cortisol on the GR and MR receptors are largely genomic.

If our above hypothesis is correct, genetic variation in the GR and MR receptors may be important sources of the genetic variation in morning cortisol levels. Since, these receptors act as *trans*acting factors (de Kloet, 2000), genetic variation in the *cis*acting elements for these activated receptors can be a second source of genetic variation.

Polymorphism(s) in the GR gene have already been associated with various aspects of cortisol metabolism such as varying basal cortisol levels (Rosmond *et al.*, 2000a; Rosmond *et al.*, 2000b) and differences in sensitivity to glucocorticoids (Huizinga *et al.*, 1998). Mutant forms of the GR gene are also found in patients with primary cortisol resistance (Ruiz *et al.*, 2001). Allelic variation in MR sensitivity is likely to further influence basal cortisol levels, although no evidence has been presented to date.

Table 10.5.

Summary Statistics of the Fit of the Genetic-Environmental Models to the four samples of cortisol.

	MODEL		**-2LL**	**df**	**Model Comparison**			
					com	**χ^2**	**df**	***p***
S1	1	Cholesky ACE sex differences, r_g DOS free	3616.172	597				
	2	Cholesky ACE sex differences, r_c DOS free	3616.153	597				
	3	Cholesky ACE sex differences	3616.153	598				
	4	Cholesky ACE no sex differences	3627.676	607	3	11.523	9	.24
	5	Cholesky AE no sex differences	3628.038	610	4	.362	3	.95
	6	Cholesky CE no sex differences	3630.672	610	4	2.996	3	.39
	7	Cholesky AE $a_{11} = a_{21}$, $a_{22} = 0$	3632.385	612	5	4.347	2	.11
S2	1	Cholesky ACE sex differences, r_g DOS free	3861.988	588				
	2	Cholesky ACE sex differences, r_c DOS free	3861.695	588				
	3	Cholesky ACE sex differences	3882.181	589				
	4	Cholesky ACE no sex differences	3870.692	598	3	8.511	9	.48
	5	Cholesky AE no sex differences	3870.922	601	4	.23	3	.97
	6	Cholesky CE no sex differences	3881.929	601	4	11.237	3	.01
	7	Cholesky AE $a_{11} = a_{21}$, $a_{22} = 0$	3882.284	603	5	11.362	2	.00

[a] The A, C and E influences are represented by a Cholesky decomposition; [b] The influence of A and C in the Cholesky decomposition are reduced to a common factor; [c] The A, C, and E influences ar represented by a common pathway model with sample specific influences; [d]A is additive genetic influences, C is shared environmental influences, E is nonshared environmental influences

Table 10.5. - continued

Summary Statistics of the Fit of the Genetic-Environmental Models to the four samples of cortisol.

	MODEL		-2LL	df	Model Comparison			
					com	χ^2	df	*p*
S3	1	Cholesky ACE sex differences, r_g DOS free	2444.214	595				
	2	Cholesky ACE sex differences, r_c DOS free	2442.230	595				
	3	Cholesky ACE sex differences	2445.157	596				
	4	Cholesky ACE no sex differences	2453.593	605	3	8.436	9	.49
	5	Cholesky AE no sex differences	2453.593	608	4	.00	3	1.00
	6	Cholesky CE no sex differences	2457.805	608	4	4.212	3	.24
	7	Cholesky AE $a_{11} = a_{21}$, $a_{22} = 0$	2456.984	610	5	3.391	2	.18
S4	1	Cholesky ACE sex differences, r_g DOS free	1337.975	560				
	2	Cholesky ACE sex differences, r_c DOS free	1336.863	560				
	3	Cholesky ACE sex differences	1338.487	561				
	4	Cholesky ACE no sex differences	1341.947	570	3	3.46	9	.94
	5	Cholesky AE no sex differences	1343.412	573	4	1.465	3	.69
	6	Cholesky CE no sex differences	1344.376	573	4	2.429	3	.49
	7	Cholesky E no sex differences	1347.354	576	4	5.407	6	.49

[a] The A, C and E influences are represented by a Cholesky decomposition; [b] The influence of A and C in the Cholesky decomposition are reduced to a common factor; [c] The A, C, and E influences are represented by a common pathway model with sample specific influences; [d] A is additive genetic influences, C is shared environmental influences, E is nonshared environmental influences

Many other sources of genetic variation should not be ruled out. These include genes that affect corticotropin-releasing factor (CRF) and adrenocorticotrophic hormone (ACTH) synthesis, the affinity, density of their receptors and their functionality. For instance, an ACTH receptor gene (Mountjoy *et al.*, 1992) has been localized on chromosome 18 (Gantz *et al.*, 1993) and mutations of this gene might lead to the disturbance of the HPAC-axis function. Further progress in understanding the genetics of individual differences in cortisol levels will be made through pharmacological and knockout studies in animals. However, although large homology probably exists between

animals and human HPAC genes, genetic linkage or candidate gene studies in humans, may be ultimately required. The high heritability of the cortisol level after awakening suggest that this may be the most useful phenotype to attempt gene finding. One of the huge advantages of a twin sample in gene finding is that observed candidate genes and unobserved genes (estimates of genetic influences through the MZ-DZ comparison) can be simultaneously tested.

Table 10.6.

Parameter estimates for additive genetic, shared environmental and nonshared environmental influences with their 95% confidence intervals.

Sample	Model	A	C	E
S11	Model 7: AE	.22 *(.09-.35)*	-	.78 *(.65-.91)*
S12	Model 7: AE	.24 *(.09-.37)*	-	.76 *(.63-.91)*
S21	Model 5: AE	.56 *(.39-.69)*	-	.44 *(.31-.61)*
S22	Model 5: AE	.59 *(.42-.72)*	**-**	.41 *(.28-.58)*
S31	Model 7: AE	.30 *(.15-.43)*	-	.70 *(.57-.85)*
S32	Model 7: AE	.21 **(.11-.30)**	**-**	.79 *(.70-.89)*
S41	Model 7: E	-	-	1.00 *(1.0-1.0)*
S42	Model 7: E	-	-	1.00 *(1.0-1.0)*

High and low basal cortisol levels in children have been associated with Externalizing and Internalizing problem behavior, respectively. From eleven studies on the association between cortisol and Externalizing behavior, nine studies report a negative association (McBurnett *et al.*, 2000; 1996; 1991; Dawes *et al.*, 1999; Van Goozen *et al.*, 1998; Scerbo and Kolko, 1994; Vanyukov *et al.*, 1993; Tennes and Krey, 1985), one study reports a positive association (Gerra *et al.*, 1998) and two studies report no difference in cortisol levels between the group with Externalizing behavioral problems and normal controls (Schulz *et al.*, 1997; Kruesli *et al.*, 1989). Children with high levels of cortisol are characterized by inhibition of temperament, higher rates of self-reported depression, parent-reported Internalizing problem behavior, social withdrawal, social anxiety, and social problems. Six of the nine reported studies on the association of Internalizing related disorders and cortisol, found a positive association (Dorn *et al.,* 1999; Granger *et al.*, 1994; Scerbo and Kolko, 1994; McBurnett *et al.*, 1991; Kagan *et al.*, 1987; Tennes *et al.*, 1986), while a negative association was found for two studies (Moss *et al.*, 1995; Vanyukov *et al.*, 1993). One study reported the finding of no association at all (Tennes and Krey, 1985).

These deviations in basal cortisol during childhood suggest a role for disturbed functioning of the hypothalamic-pituitary adrenocortico (HPAC)-axis in the pathogenesis of these behavioral disorders. Some circumstantial evidence for this chain of events exists as significant genetic effects have also been found on problem behavior. For instance, van der Valk observed both genetic effects on problem behavior in a large sample of 3-year old twins (van der Valk et al., 1998[a]), as well as stable genetic influences on problem behavior at the age of 10 and 15 years in biologically related and unrelated adoptees (van der Valk *et al.*, 1998[b]). Significant influences of genetic factors have been found on stability in problem behavior in large longitudinal sample of Dutch twins followed from age 3 to age 12 (Bartels *et al.*, 2002[e]). Comparable results of genetic influences on stability have been found in other studies (Verhulst & van der Ende, 1993; van den Oord, 1994; Koot, 1995; Edelbrock *et al.*, 1995; Schmitz *et al.*, 1995). Although these findings allow a scenario in which genetic influences on HPAC-axis functioning translate to a genetic risk for behavioral problems, the opposite – an effect of genetically determined problem behavior on HPAC-axis functioning - cannot be ruled out. Also, an underlying genetic defect may cause problem behavior as well as deviant HPAC function without a direct causal link between these two (pleiotropy) (Blizard, 1992). Future multivariate modeling of cortisol and problem behavior could resolve this main issue of causality.

Finally, it should be emphasized that cortisol in these children was collected at home. Hence, the results of this study give a good insight in the genetic and environmental influences on basal cortisol levels but not on cortisol levels in response to specific physical or emotional stressors. Individual differences in stress reactivity may well be a key factor in the link between HPAC-axis functioning and behavior. Cortisol reactivity to standardized stressors, therefore, would be served by future examination in genetically related children.

11

Cortisol, Behavioral Problems and Cognition

INTRODUCTION

Cortisol and childhood psychopathology

Disturbances in hypothalamic-pituitary-adrenal (HPA) regulation, possibly due to prenatal exposure to glucocorticoids causing prenatal programming, are associated with affective and anxiety disorders in humans (Lopez *et al.*, 1998; Holsboer and Barden, 1996). Many studies report an association between cortisol levels and Externalizing and Internalizing problem behaviors during childhood (for an overview see Table 11.1). Previous studies report on relations between low levels of cortisol and aggressive and antisocial behavior. From the eleven studies on the association between cortisol and Externalizing behavior, presented in Table 11.1, nine studies report a negative association (McBurnett *et al.*, 2000; 1996; 1991; Dawes *et al.*, 1999; Van Goozen *et al.*, 1998; Scerbo and Kolko, 1994; Vanyukov *et al.*, 1993; Tennes and Krey, 1985), one study reports a positive association (Gerra *et al.*, 1998) and two studies report no difference in cortisol levels between the group with Externalizing behavioral problems and normal controls (Schulz *et al.*, 1997; Kruesli *et al.*,1989). In contrast, children with high levels of cortisol are characterized by inhibition of temperament, higher rates of self-reported depression, parent-reported Internalizing problem behavior, social withdrawal, social anxiety, and social problems. Six of the nine reported studies on the association of Internalizing related disorders and cortisol, found a positive association (Dorn *et al.,* 1999; Granger *et al.*, 1994; Scerbo and Kolko, 1994; McBurnett *et al.*, 1991; Kagan *et al.*, 1987; Tennes *et al.*, 1986), while a negative association was found for two studies (Moss *et al.*, 1995; Vanyukov *et al.*, 1993). One study reported the finding of no association at all (Tennes and Krey, 1985).

Different measurement methods for both cortisol and psychopathology may partly explain the inconsistency in the results of previous studies. One difference across studies involves the hormone collection procedures, such as sampling from serum, urine or saliva, taking single versus multiple measurements and taking samples in response to a stressor or without a stressor to define basal daily variation. All methods of sample collection have their pro's and cons and it depends on the aim of the studies which method is more appropriate (Aardal and Holm 1995; Kirschbaum and Hellhammer 1994; Trainer *et al.*, 1993; Kirschbaum and Hellhammer, 1989; Riad-Fahmy *et al.*, 1982). In general, saliva collection is the most practical and stress-free method of cortisol collection in a large group of subjects (both adults and children).

The reason why blood and urine sampling have been used more often, is probably historical as the development of the "Salivette" has taken place fairly recently and the knowledge about the use of saliva as a representative biological fluid has increased over

the past years. Further, both blood and saliva can provide information on the diurnal rhythm, while urine measures represent the cortisol production over a period of time.

Additionally, in measuring basal levels of cortisol multiple sampling is essential according to the circadian rhythm. In the studies mentioned in Table 11.1, a huge variation in number of samples and time of sampling can be observed creating possible differences in results of distinct studies. Cortisol shows a strong circadian rhythm, but the secretion of cortisol is also a classical endocrine response to stress. In order to compare studies on cortisol levels a clear distinction should be made between basal cortisol levels and cortisol reactivity to stressors or chemical stimuli (e.g. ACTH), because these cortisol levels are based on a completely different physiological mechanism.

Other differences across studies on cortisol and problem behavior can rise due to determination of the aspects of problem behavior and strength of problem behavior under study. Validity in behavioral data have been diverse as represented by observations, ratings by professional or parental or teacher questionnaires. Additionally, some studies focus on a single dimension of problem behavior, such as aggression while other studies take multiple dimensions into account, like Internalizing or Externalizing behavior. Another problem with studying problem behavior is the comorbidity and definition of the problem behavior syndromes, especially in clinical samples.

The inconsistent result could also partly be explained by the samples used. Most studies use small clinical samples and it had yet to be established if the results found in clinical samples can be generalized to the general population. Further, the power of a study to detect an association depends partially on the sample size.

Finally, in none of the studies data are available to gain insight into the physiological mechanism underlying the association between problem behavior and cortisol levels. Most studies focus on the presence or absence of an association solely. A single study considered the impact of glucocorticoids (GC, cortisol) on cognitive and behavioral development in children who were exposed to repeated antenatal GC treatment but not born before term (Trautman *et al.*, 1995). Children exposed to dexamethasone in early pregnancy, because of increased risk of congenital adrenal hyperplasia, showed significant increases in emotionality, unsociability, avoidance, and behavioral problems. Higher cortisol reactivity in young children following a parent-child conflict task may contribute to subsequent Internalizing symptoms over a 6-month period, suggesting effects on general cognitive and emotional performance that might have implications for later psychopathology (Granger *et al.*, 1996). Further, studies are required to establish the impact of elevated GC, due to repeated antenatal GC treatment or maternal stress, on behavioral outcome in children just after birth and later in life.

Table 11.1. *Overview of studies on cortisol and childhood psychopathology*

EXTERNALIZING PROBLEM BEHAVIOR				
Study	*Subjects*	*Cortisol measurements*	*Behavior measurement*	*Results / direction of association*
McBurnett *et al.* (2000)	- 38 clinic-referred boys (m.a.9.75[b])	- single saliva sample in year 2 and 4 of the study - time of sampling varied across the day	- The NIMH Diagnostic Schedule for Children, Version 2 - Peer nomination counts	- cortisol levels strongly and inversely related to aggressive CD, peer aggression nomination and ODD (-) [a] - cortisol levels positively associated with age of first appearance of aggressive behavior (+)
Dawes *et al.* (1999)	- 150 high substance abuse risk boys (age 10-12 years) - 147 low substance abuse risk boys (age 10-12 years)	- single saliva sample just before (9.00 AM) ERP - single saliva sample just after (10.15 AM) ERP	- Behavior self-regulation scale - Dimensions of Temperament Survey - The Disruptive Behavior Disorder Rating Scale - Teacher Report Form	- basal cortisol and cortisol reactivity lower in high risk group (-) - negative correlation between cortisol and number of conduct disorder symptoms in high risk group (-)
Van Goozen *et al.* (1998)	- 21 oppositional-defiant disorder boys (m.a. 10.2 years) - 31 normal controls (m.g. 9.6.years)	- six saliva samples between 9.00 AM and 11.30 AM; two samples during baseline registration, three samples during stress, one sample nonstress.	- ODD; extensive psychiatric assessments - Child Assessment Schedule - Child Behavior Checklist - Teacher Report Form - Wechslers Intelligence Scale for Children - Stress was induced by frustration, provocation and aggression	- cortisol levels in ODD group overall lower than NC (-) - effect of stress minimal on both groups
Gerra *et al.* (1998)	- 30 male peripubertal adolescents (m.a. 12.7 years)	- blood for hormone assays was drawn 30 minutes after inserting catheter (time 0) - second blood sample after 30 minutes and three stress tests	- Aggression measured by trained psychologists - Children Personality Questionnaire - Children's Depression Inventory - Test Anxiety Inventory - Stress tests; Stroop, mental arithmetic test, public speaking	- basal levels of cortisol higher in high-normal aggressiveness group (+)

Table 11.1 - continued. *Overview of studies on cortisol and childhood psychopathology*

EXTERNALIZING PROBLEM BEHAVIOR				
Study	*Subjects*	*Cortisol measurements*	*Behavior measurement*	*Results / direction of association*
Schulz *et al.* (1997)	- 23 aggressive boys (m.a. 9.0 years) - 27 nonaggressive boys (m.a. 9.0 years) - all subjects met DSM-III-R criteria for ADHD	- catheter in forearm vein at 8.00 AM - samples of blood for cortisol obtained at 9.45 AM and 9.55 AM	- IOWA Teacher's Questionnaire - Child Behavior Checklist - Diagnostic Interview Schedule for Children - Wechslers Intelligence Scale for Children	- no difference in cortisol levels between aggressive and nonaggressive groups (0) - no difference in cortisol levels when subjects were divided by presence of absence of CD (0)
McBurnett *et al.* (1996)	- 67 boys (m.a. 9.6 years)	- single saliva sample in year 2 and 4 of the study - time of sampling varied across the day	- Diagnostic Interview Schedule for Children	- total number of aggressive symptoms inversely associated with salivary cortisol (-) - age at first aggressive CD symptom associated with salivary cortisol in year 2 - total number of CD symptoms over 4 years inversely associated with cortisol from year 2 (-)
Scerbo and Kolko (1994)	- 40 clinic-referred disruptive children (37 boys and 3 girls) (m.a. 11.18 years) - 19 diagnosed with CD - 23 diagnosed with ADHD (11 comorbid CD; 8 comorbid ODD) - 17 diagnosed ODD	- single saliva sample between 9.00 and 10.00 AM after 30-minute quiet nonstress period, during the first 8 days of the program	- standardized clinical interview based on the Schedule for Affective Disorders and Schizophrenia for School-Age Children - Child Behavior Checklist - Teacher Report Form - Overt Aggression Scale - Iowa-Conners Questionnaire - Conners Parent Questionnaire	- cortisol negatively correlated with staff-rated inattention/overactivity (-) - cortisol negatively correlated with staff-rated oppositional behavior (-)

Table 11.1 - continued. *Overview of studies on cortisol and childhood psychopathology*

EXTERNALIZING PROBLEM BEHAVIOR				
Study	*Subjects*	*Cortisol measurements*	*Behavior measurement*	*Results / direction of association*
Vanyukov *et al.* (1993)	- 78 sons of fathers who have a diagnosis of psychoactive substance use disorder after age 18 (m.a. 10.9 years) - 72 Normal controls; sons of fathers who have no psychiatric disorder (m.a. 10.9 years)	- single saliva sample collected at 9.00 AM during a period of rest, while paste electrodes were placed on the child's scalp	- expanded version of the Kiddie-Schedule for Affective Disorders and Schizophrenia	- no overall differences in resting cortisol concentration between the substance abuse and control groups (0) - cortisol levels lower in sons of those fathers diagnosed CD before age 18 and anti social personality (ASP) as adult (-) - cortisol concentration negative correlated with the child' s CD symptom count (-) - cortisol concentration negatively associated with ASP symptom count of the father (-)
McBurnett *et al.* (1991)	- 67 clinic-referred boys divided in; 7 CD, no anxiety (m.a. 9.6); 11 CD, anxiety (m.a. 9.3); 28 no CD, no anxiety (m.a. 9.3); 21 no CD, anxiety (m.a. 9.0)	- single saliva sample before psychological assessment - time of collection was recorded	- children grouped to DSM-III-R diagnosis of CD and anxiety disorder	- boys with and without anxiety disorder differed in salivary cortisol only when comorbid CD was present
Kruesli *et al.* (1989)	- 19 boys (15 ADDH, 14 CD, 4 ODD) (m.a. 125.5 months) - 18 normal controls (m.a. 125.7 months)	- 24-hour urine collection (9.00 AM to 9.00 AM next day)	- Conners Teacher Rating Scale - DICA - DICA-P	- no difference in urinary free cortisol between patients with disruptive behavior diagnosis and normal controls (0)
Tennes and Kreye (1985)	- 70 second graders (38 boys, 32 girls) (m.a. 7.7)	- once a month urine samples arriving at school and two hours later (11.00 AM) - on days with achievement tests	- 2 hour observation on day of urine collection - Test Anxiety Scale for Children (TASC) - Several intelligent tests - Teacher ratings on behavior	- cortisol production higher on test days - problem behavior (aggressive behavior) negative correlated with cortisol (-)

Table 11.1. – continued. *Overview of studies on cortisol and childhood psychopathology*

INTERNALIZING PROBLEM BEHAVIOR				
Study	*Subjects*	*Cortisol measurements*	*Behavior measurement*	*Results*
Dorn *et al.* (1999)	- 9 children with premature adrenarche (8 girls, 1 boy; aged 7 to 11 years) - 21 children with on time adrenarche (8girls, 12 boys) (m.a. 8.0)	- single saliva sample at home (±noon) at day of testing (T_h) - 3 saliva and 3 blood samples (starting 1.30 PM, 20 minutes later, 40 minutes later; T_0, T_{20}, T_{40})	- The Diagnostic Interview Schedule for Children - Child Behavior Checklist - Children's Depression Inventory - State Trait Anxiety Inventory for Children - Wechslers Intelligence Scale for Children-R	- PA more self-reported depression, more behavior problems, lower cognitive functioning - PA group higher saliva cortisol for T_h, T_0, T_{20}, T_{40} (+) - PA group higher serum cortisol for T_0 (+)
Moss *et al.* (1995)	- 81 high substance abuse risk boys (m.a. 10.86) - 103 low substance abuse risk boys (m.a. 11.1)	- saliva collection began at 9.00 AM - saliva samples before and after ERP (interval 35 minutes)	- Kiddie-SADS-E (presence of psychiatric disorders) - Spielberger State-Trait Anxiety Scale for Children - Child Behavior Checklist - Teacher Report Form	- high risk group had lower pre-and post-ERP salivary cortisol concentration than low risk group (-) - high risk group show reduced cortisol responsivity than low risk group (-)
Granger *et al.* (1994)	- 102 clinic-referred children (62 boys and 40 girls) (m.a. 12.1 years)	- saliva samples collected before and 20 minutes after parent-child interaction task - average time for pretask and posttask saliva collection were 4.26 PM and 5.27 PM (mean interval 1 hr)	- Parent-Child Conflict Discussion Task was used as the psychologically challenging event - Youth Self-Report - Child Behavior Checklist - Social Anxiety Scale for Children - Children's Depression Inventory	- posttask cortisol levels correlated with social withdrawal and social anxiety (+)

Table 11.1 - continued. *Overview of studies on cortisol and childhood psychopathology*

INTERNALIZING PROBLEM BEHAVIOR				
Study	*Study*	*Study*	*Study*	*Study*
Scerbo and Kolko (1994)	- 40 clinic-referred disruptive children (37 boys and 3 girls) (m.a. 11.18 years) - 19 diagnosed with CD - 23 diagnosed with ADHD (11 comorbid CD; 8 comorbid ODD) - 17 diagnosed ODD	- single saliva sample between 9.00 and 10.00 AM after 30-minute quiet nonstress period, during the first 8 days of the program	- standardized clinical interview based on the Schedule for Affective Disorders and Schizophrenia for School-Age Children - Child BehaviorChecklist - Teacher Report Form - Overt Aggression Scale - Iowa-Conners Questionnaire - Conners Parent Questionnaire	- cortisol positively correlated with parent-rated internalizing behavior (+) - cortisol positively correlated with parent-rated depression (+)
Vanyukov *et al.* (1993)	- 78 sons of fathers who have a diagnosis of psychoactive substance use disorder after age 18 (m.a. 10.9 years) - 72 Normal controls; sons of fathers who have no psychiatric disorder (m.a 10.9 years)	- single saliva sample collected at 9.00 AM during a period of rest, while paste electrodes were placed on the child's scalp	- expanded version of the Kiddie-Schedule for Affective Disorders and Schizophrenia	- no overall differences in resting cortisol concentration between the substance abuse and control groups - cortisol concentration negatively associated with ASP symptom count of the father (-)
McBurnett *et al.* (1991)	- 67 clinic-referred boys divided in: 7 CD, no anxiety (m.a. 9.6); 11 D, anxiety (m.a. 9.3); 28 no CD, no anxiety (m.a. 9.3); 21 no CD, anxiety (m.a. 9.0)	- single saliva sample before psychological assessment - time of collection was recorded	- children grouped to DSM-III-R diagnosis of CD and anxiety disorder	- association between cortisol and anxiety disorder (+) - normalized cortisol concentrations correlated with the number of overanxious and separation anxiety symptoms separately and with the total number of symptoms of both types of anxiety among boys with CD (+)

Table 11.1 cont. *Overview of studies on cortisol and childhood psychopathology*

INTERNALIZING PROBLEM BEHAVIOR				
Study	*Subjects*	*Cortisol measurements*	*Behavior measurement*	*Results*
Kagan *et al.* (1987)	- 60 consistently inhibited children (equal boys and girls) - 60 consistently uninhibited children (equal boys and girls) - all aged 21 or 31 months	- saliva samples at 21 months and after 5.5 years - before and after laboratory session - at home during 3 days in early morning hours	- Observation	- inhibited children higher cortisol levels than uninhibited children in both home and laboratory (+)
Tennes *et al.* (1986)	- 30 second graders (16 boys, 14 girls)	- urine samples collected between 9.00 AM and 11.00 AM	- 2 hour observation on day of urine collection - Teacher ratings on behavior - Cognitive Abilities Test	- mean cortisol levels on test days were elevated over cortisol on normal days - cortisol positively correlated with social affiliative behavior (+)
Tennes and Kreye (1985)	- 70 second graders (38 boys, 32 girls) (m.a. 7.7)	- once a month urine samples arriving at school and two hours later (11.00 AM) - on days with achievement tests	- 2 hour observation on day of urine collection - Test Anxiety Scale for Children (TASC) - Several intelligent tests - Teacher ratings on behavior	- cortisol production higher on test days - no significant between TASC and cortisol (0) - moderate and high IQ higher cortisol on test days - social competence and task orientation positively correlated with cortisol

[a] (+) = positive relation between cortisol levels and problem behavior, (-) = negative relation between cortisol levels and problem behavior, (0) = no relation between cortisol and problem behavior; [b] m.a. = mean age

To investigate the nature of the association between cortisol and behavioral problems insight into the etiology of individual differences in basal cortisol levels in children is essential.

Thus, while some studies suggest an association between cortisol levels and common childhood psychopathology, there have been inconsistent findings, which may vary as a function of the reliability of cortisol sampling, problem behavior ratings, the use of clinical or nonclinical samples, and the sample size used. The purpose of this chapter is to assess the relation between baseline salivary cortisol and Internalizing and Externalizing behavior in a large non-clinical sample. The consideration of this normally developing population can be informative, because clinic-referred children's problems may not represent extremes on behavioral, cognitive, and emotional dimensions. Therefore studying a sample of the general population is useful to gain insight into healthy functioning children.

Cortisol and cognition

An association between cortisol and cognition can exist because glucocorticoid receptors are expressed in parts of various brain regions, so glucocorticoids are involved in the regulation of neural metabolism, physiologic functions, and gene expression in the brain, particularly in the hippocampus. A range of evidence supports the role of the hippocampus in declarative memory performance (e.g. Squire, 1992; Monk and Nelson, 2002) and high levels of cortisol seem to be particularly damaging to the hippocampus (Lupien *et al.*, 1998). Some evidence supports the notion that high circulating levels of cortisol are correlated with impaired psychological performance from childhood though to adult life. For instance, case-control study designs indicate decreased memory performance during corticosteroid treatment of asthmatic children (Bender *et al.*, 1988) and decreased verbal declarative memory in corticosteroid-treated patients vs. matched medical control subjects (Keenan *et al.*, 1996). Investigators also reported inverse correlations between memory performance and plasma concentrations of cortisol in patients with Cushing syndrome, dementia of the Alzheimer type, schizophrenia, and depression (Whelan *et al.*, 1980; Starkman *et al.*, 1981;; Rubinow *et al.*, 1984 Heuser *et al.*, 1988; Newcomer *et al.*, 1998). Evidence for an inverse relation between several days of exposure to cortisol and verbal declarative memory in healthy subjects is presented in a study by Newcomer and colleagues (1999). Based on these reported results it is important to gain insight into the relationship between basal cortisol levels and cognitive functioning in the normal population, especially in children.

METHODS

Participants and Analysis

Internalizing and Externalizing behavior problems are rated by mothers and fathers at age 12, using the Child Behavior Checklist in a large sample of twins (For details see chapter 6). Maternal ratings were available for 1,481 twin pairs and paternal ratings were available for 1,156 twin pairs. In a subsample data on cognition are collected at age 12, using the WISC-R (for details see chapter 7) as well. A measure of Full Scale IQ is available for 381 children. Bivariate Pearson's correlations were calculated for boys and girls separately using SPSS/windows 10.

Table 11.2.

Pearson correlations between Cortisol and Internalizing behavior, Externalizing behavior, and cognition.

		Cortisol 1[a]	Cortisol 2	Cortisol 3	Cortisol 4
Int12$_m$	♂	.111 (163)	.068 (161)	-.119 (160)	.154 (157)
	♀	.010 (171)	.135 (172)	.052 (171)	.056 (164)
Int12$_f$	♂	.117 (150)	-.041 (150)	-.129 (147)	-.025 (144)
	♀	.112 (154)	-.010 (155)	-.005 (156)	.037 (151)
Ext12$_m$	♂	.034 (168)	.025 (166)	-.078 (165)	.132 (161)
	♀	.012 (174)	.286 (175)**	.166 (174)	.161 (167)*
Ext12$_f$	♂	-.103 (151)	-.104 (151)	-.175 (148)*	-.011 (145)
	♀	.025 (155)	.090 (156)	.147 (157)	.007 (152)
IQ12	♂	-.050 (177)	-.006 (175)	-.041 (174)	-.068 (170)
	♀	-.085 (178)	-.040 (179)	-.032 (178)	.011 (171)

[a] the mean of the same samples at the two consecutive days; 1= first sample (before getting up); 2=second sample (half an hour after awakening); 3=third sample (before lunch); 4=fourth sample (evening); [b] number of children used to calculate this correlation; [c] *correlation is significant a .05 level; ** correlation is significant at .01 level.

RESULTS

As a preliminary finding to support future research, correlations between cortisol levels and childhood psychopathology or cognition are calculated. Table 11.2 presents the correlations between basal cortisol levels and Internalizing behavior, Externalizing behavior, and cognitive abilities. The results presented in Table 11.2 show no overwhelming evidence for the significance of a simple association between the traits under investigation. Further the significant correlations that are found point more into the direction of a change finding than a consistent pattern of associations. One of the most striking findings is that two of the three significant correlations found for Externalizing behavior and cortisol are positive, while most previous studies report a negative association.

Because the use of the broadband scales Internalizing and Externalizing may overshadow an association between certain kinds of problem behavior and cortisol, correlations were calculated for anxiety and aggressive behavior as well. In Table 12.3, the picture that emerges from Table 12.2 is confirmed. For cortisol and anxiety and aggressive behavior no significant or consistent associations are found. Further, the significant correlations are in the opposite direction than expected, with higher cortisol levels being assessed with higher levels of aggression.

The results of these data do not completely rule out the possibility of an association between problem behavior and cortisol levels in children. For instance, sampling at a normal day during the week at home reflects the steady, basal activity of the adrenal cortex of an organism, while measuring cortisol before and after a stressor reflects the organism's adrenocortical reactivity to its environment (Wolff *et al.*, 1964). Cortisol secretion has been found to rise in stressful or anxiety-provoking situations, although there is considerable variability in this rise between individuals (Mason, 1968). Therefore, it is possible that there is a relationship between childhood problem behavior and cortisol levels in response to a stressor, while in this study only basal cortisol levels have been used.

The findings of these preliminary analyses may also be a result of lack of power. The power to detect a significant association depends on a number of factors, namely, the true size of the effect in question, the probability level chosen and the sample size (Cohen, 1992). The number of participants in which daytime cortisol was measured was high in comparison to previous studies in the field. However, when only moderate correlations are found in clinical samples, the effect in non-clinical referred children might be even smaller and in that respect larger samples are necessary to find this association. Finally, for the determination of basal cortisol levels only four sample of saliva are collected a day. As mentioned in Chapter 9, probably more than four samples a day are necessary to get a reliable picture of basal cortisol levels in children.

Table 12.3.

Pearson correlations between Cortisol and Anxiety and Aggressive behavior.

		Cortisol 1[a]	Cortisol 2	Cortisol 3	Cortisol 4
Anx12$_m$	♂	.182 (170)*	.144 (168)	-.069 (167)	.080 (163)
	♀	-.067 (180)	.034 (181)	.006 (180)	.071 (173)
Anx12$_f$	♂	.058 (152)	.040 (152)	-.103 (149)	-.004 (146)
	♀	.049 (160)	.024 (161)	.031 (162)	-.050 (157)
Agg12$_m$	♂	.100 (169)	.063 (167)	.004 (166)	.134 (162)
	♀	-.016 (179)	.200 (180)**	.102 (179)	.205 (172)**
Agg12$_f$	♂	-.045 (152)	-.061 (152)	-.199 (149)*	-.016 (146)
	♀	.026 (160)	.097 (161)	.140 (162)	.019 (157)

[a] the mean of the same samples at the two consecutive days; 1= first sample (before getting up); 2=second sample (half an hour after awakening); 3=third sample (before lunch); 4=fourth sample (evening); [b] number of children used to calculate this correlation; [c] *correlation is significant a .05 level; ** correlation is significant at .01 level.

FUTURE DIRECTIONS

Based on the contrast of our findings with findings from previous studies some future directions in this kind of research should be considered. First, our assessment instrument for problem behavior, the Child Behavior Checklist, does not measure DSM or ICD psychiatric diagnoses used in most previous studies. Using the CBCL results in a continuous variable of problem behavior. To overcome the discrepancy in the diagnosis of psychopathology between this study and the previous studies it is possible to select children from the population sample. Selection criteria can be based on CBCL scores in the range that these children are likely to meet DSM criteria for Internalizing or Externalizing behavior. In the future we are also able to create a large dataset with children who show persistent problem behavior from age 3 throughout development. Children at the extreme ends of the distribution can be selected based on the presence or absence of persistence in problem behavior.

Some other aspects of the cortisol data can be useful to investigate. In general cortisol shows a well-documented circadian rhythm resulting from stimulation of paraventricular neurons by pacemaker cells located in the supriciasmatic nucleus. Peak cortisol levels are observed shortly after awakening with steadily decreasing values thereafter. The trough of cortisol secretion is reached around midnight with only minimal levels of cortisol detectable (Anders, 1982; Desir *et al.*, 1980; Weitzman *et al.,* 1971). In two recent studies, though, variabilities are found in this circadian rhythm which was believed to be relatively homogeneous. Smyth and colleagues (1997) collected six saliva samples at random points throughout the day for two consecutive days in a group of community-dwelling individuals. 51% of their sample showed strong, decreasing patterns of cortisol on both days, 31% showed inconsistent cycles where one day was typical and the other flattened, and 17% showed flattened cycles on both days. This result is replicated in a study by Stone and colleagues (2001), who found that 15% of community individuals did not show the typical diurnal rhythm. In respect to investigating a possible association between cortisol levels and childhood psychopathology or cognitive abilities it might me useful to first further investigate the circadian rhythm in adults as well as children. If variation in cycles is found it is interesting to see whether splitting the sample in three groups, (I) always rising, (II) inconsistent pattern, (III) always flattened, results in distinct levels of problem behavior between the groups. However, by doing so, we must take into consideration that little is known about reproducibility of diurnal rhythms over longer periods of time, namely, months or years.

In conclusion, the often cited association between problem behavior and cortisol levels in clinical referred groups or small samples was not paralleled by an association between Internalizing behavior, Externalizing behavior and cortisol levels in a population sample. Future research is necessary to replicate or falsify this finding.

BACKGROUND HYPOTHESIS

Why is an association between cortisol, psychopathology and cognition expected?

A possible moderator for individual differences in cognition and problem behavior is cortisol. In other words, individual differences in psychopathology or cognitive abilities could be a results of individual differences in cortisol levels as a result of genetic and environmental influences. In this matter cortisol can be considered as an endophenotype. In general endophenotypes represent biological, neurophysiological, electrophysiological, and behavioral indices of the pathways that connect genes and the trait under investigation. The possible moderator effects of cortisol are partly based on a phenomenon dubbed 'fetal programming' (for a review see: Welberg and Seckl, 2001; Matthews, 2000). In the prenatal period the development of the brain is influenced by

hormones, secreted by the pituitary and the gonads (Collaer and Hines, 1995; Edwards *et al.*, 1993; Sikich and Todd, 1988). These early effects are referred to as 'programming effects' because of their possible influence on the development and structure of the brain.

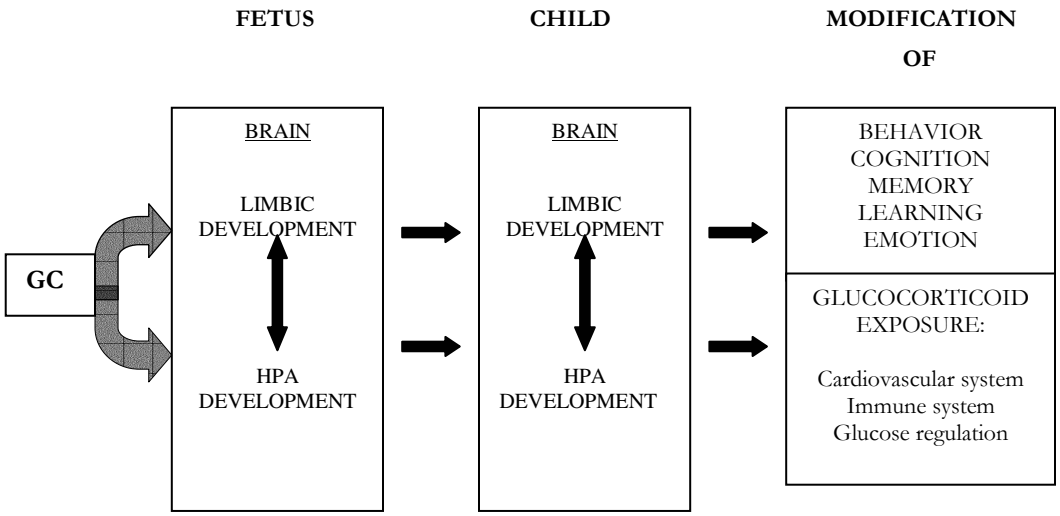


Figure 11.1.

Diagrammatic representation of the routes by which prenatal GC exposures programs postnatal behavior and neuroendocrine function. The fetal limbic system (primarily the hippocampus), hypothalamus, and anterior pituitary express high concentrations of corticoid receptors, and are sensitive to GCs. Exposure to exogenous GC at this time alter development and subsequent function of both the limbic system and the HPA axis. The hippocampus regulates HPA function, and endogenous GCs (the end product of HPA activation) modify many aspects of limbic function. In the periphery, the overall effect of programming during development will be altered exposure to endogenous GC throughout life. Increased exposure will predispose to a number of neurologic, metabolic and cardiovascular diseases, whereas reduced exposure may act to protect against these diseases (Matthews, 2000).

Glucocorticoids (GC) (e.g. cortisol) are essential for normal development. However, excess exposure has deleterious effects, inhibiting fetal growth and altering the trajectory of tissue maturation. The hypothalamic-pituitary-adrenal axis (HPA-axis), which is the central to the integration of the individual's endocrine and behavioral response to stress, appears highly sensitive to excess GC exposure during development (Figure 11.1).

For instance, exposure of pregnant rats to exogenous or endogenous glucocorticoids not only reduces birth weight, produces permanent hypertension, hyperglycemia, and

hyperinsulinia, but also results in behavioral inhibition, impaired coping and adult affective dysfunction (Welberg *et al.*, 2001; Nyirenda *et al.*, 1998; Levitt *et al.*, 1996; Lindsay *et al.*, 1996a and 1996b). In humans, increased exposure to glucocorticoids (cortisol) can occur due to maternal stress during pregnancy. Several behavioral abnormalities have been reported in children exposed to 'prenatal stress' (Stott, 1973).

Besides these prenatal influences of hormones, persistently higher postnatal cortisol levels, from whatever cause, might endanger the functional integrity of the brain and hence the probability of cognitive dysfunction and subsequent psychopathology. The developmental origins of these risks remain unclear but early environmental adversities are one set of candidates. For instance, evidence suggests that early adverse experiences, like childhood abuse or parental separation, play a prominent role in development of mood and anxiety disorders and that corticotrophin-releasing hormone (CRH) systems may mediate this association (Mullen *et al.*, 1996; Heim *et al.*, 2000). Further evidence for this association has been assembled in animal models, where prenatal and early developmental stress, often related to parental rearing, have been shown to cause long-lasting or even permanent alteration of the HPA axis (Plotsky and Meaney, 1993; Levine, 1994; Schmidt *et al.*, 2002). Not only early experiences, but also experiences later in life can influence HPA axis activity. For example, trauma survivors with posttraumatic stress disorder such as Vietnam veterans, holocaust survivors or victims of abuse are characterized by decreased urinary cortisol level as compared to healthy controls (see, among others, Yehuda *et al.*, 1991; 1995, 2000). Accordingly, environmental challenges are important in the development of HPA axis disregulation and stress-related diseases.

Research on the functional effects of pre- and postnatal cortisol levels in children and adolescent is complicated by developmental changes in cognitive abilities, childhood psychopathology and hormonal milieu. So, insight into the developmental process of genetic and environmental influences on cognition and Internalizing and Externalizing problem behavior and knowledge on the cause of individual differences in basal cortisol levels is essential to investigate a possible relationship between these two variables. Cortisol levels at age 12, as used in this project, may or may not reflect prenatal cortisol levels, which are of significance in the fetal programming hypothesis. One of the ways to gain more insight into this complex system of pre- and postnatal cortisol levels is to collect umbilical cord blood from twins. Cortisol levels determined at birth may reflect the prenatal cortisol levels. Ideally, these twins will be followed from birth onwards to investigate the genetic and environmental influences on the developmental pattern of cortisol levels, behavioral problems and cognition.

12

Discussion

This final chapter gives an overall conclusion followed by a summary of the empirical results presented in this thesis. The implications for future research are discussed.

Overall conclusion

Based on the results of this thesis it can be concluded that genetic influences are pervasive during childhood. Genes affect two major phenotypes, behavioral problems and cognitive abilities, as well as a possible endophenotype, basal cortisol levels. Genes also affect various aspects of development and are very important for continuity and change throughout development.

It can be stated that genetic studies are not only important for estimating size of genetic and environmental effects, but can shed light on fundamental questions in child development. For this reason, genetic studies supply phenotypic studies and are an essential addendum for research in developmental psychology.

Childhood psychopathology

Prevalence

Behavioral and emotional problems are common among children. Twenty-five to 30% of the children in the twin sample showed behavioral problem (Table 12.1).

Table 12.1.

Prevalence of behavioral problems in boys and girls according to mother or father ratings (T-score ≥60)

	♂	♀
age 3	31.8%	27.9%
age 7	31.2%	25.2%
age 10	30.0%	24.0%
age 12	28.8%	25.2%

To be more specific the prevalence of behavioral problems varies depending on the kind of behavior, the rater, and the age and gender of the child (Table 12.2). Internalizing problems (anxious/depressed behavior, withdrawn behavior) are more prevalent in girls than boys and Externalizing problems (aggressive behavior, rule breaking behavior) are more prevalent in boys than girls. Overall, mothers report more problem behaviors than fathers. Finally, prevalence of problem behavior, as rated by the parents, decreases over age.

For both the Internalizing and the Externalizing scale, the ratings given to the twins were quite similar to the ratings in the norm sample (Verhulst *et al.*, 1996). In previous

studies, using an overlapping sample, comparable levels of problem behavior were found (Van der Valk *et al.*, 1998, 2001). Since the prevalence and means of problem behaviors in the twin sample are comparable to children of the Dutch population, generalization of the results of this twin study to Dutch children of the same age seems allowed.

Table 12.2.

Prevalence of Internalizing, Externalizing or both kinds of problem behavior in the longitudinal twin sample based on mother and father ratings (T-score ≥ 60).

	Mother ratings		Father ratings	
	♂	♀	♂	♀
Int3[a]	8.5%	10.2%	8.4%	9.0%
Int7	5.3%	9.4%	6.0%	8.7%
Int10	5.5%	9.1%	4.8%	7.8%
Int12	5.6%	9.0%	4.7%	9.2%
Ext3[b]	11.1%	7.9%	11.2%	7.2%
Ext7	11.5%	4.6%	12.3%	5.3%
Ext10	11.2%	4.0%	11.7%	5.2%
Ext12	11.2%	4.3%	12.5%	5.2%
Int3 & Ext3[c]	7.3%	5.7%	6.1%	5.2%
Int7 & Ext7	6.3%	4.8%	6.9%	5.0%
Int10 & Ext10	6.8%	4.5%	6.8%	4.7%
Int12 & Ext12	5.4%	4.9%	7.0%	3.7%

[a] Internalizing problems only; [b] Externalizing problems only; [c] Internalizing and Externalizing problems

Development of Internalizing and Externalizing behavior

A main objective of this thesis was to investigate the etiology of individual differences in the development of psychopathology. For the study of development of Internalizing and Externalizing problem behavior parental ratings on psychopathology collected at ages 3, 7, 10, and 12 years in birth cohorts 1986 up to and including 1993 were used.

To gain a first insight into the development of Internalizing and Externalizing problem behavior structural equation modeling techniques were used in a large longitudinal sample of Dutch twins, in which the mother of the twins rated behavioral problems. The observed stability coefficients for the four- seven-, and nine-year time intervals were, respectively, .37, .33, .30 for Internalizing behavior and .55, .49, and .48 for Externalizing behavior. The phenotypic correlations were in line with results from previous large scale

longitudinal studies (Verhulst and Van der Ende, 1992a; 1992b; Ghodsian *et al.*, 1980; Richman *et al.*, 1982; Graham and Rutter, 1973).

Besides a common factor structure for the influences of shared environment on the development of behavioral problems, a simplex structure for genetic influences on the developmental process was found. Based on these developmental patterns it was estimated that stability in Internalizing problem behavior is for 43% accounted for by additive genetic factors and for 47% by shared environmental factors. For Externalizing behavior, stability was represented by additive genetic transmission factors explaining 67% of the stability over the years, on average, for boys and 53% of the stability over the years, on average, for girls. Stability was further accounted for by shared environmental influences, explaining 27% and 40% of the total stability for boys and girls, respectively. Change in Internalizing and Externalizing behavior in both boys and girls could be mainly explained by nonshared environmental influences. Genetic innovations and age specific shared environmental influence also accounted for some change in both problem behaviors over the years.

Our finding of different developmental patterns for the distinct sources of variance (additive genetic and shared environmental) has important implication for the prevention of later maladjustment. The shared environmental influences, for instance, exert a continuous influence from their time of onset. So, children who continue to experience adverse shared environment are at risk for later maladjustment. For additive genetic influences, parts of previous effects are transmitted to later ages. However, genetic influence is less static due to new genetic influences that come into play at each age. Nonshared environmental influences seem to be important for age-specific behavior problems and have almost no developmental significance. This implies that influences of nonshared environment are important but that they are mostly of transient nature and specific to a specific moment in time.

In studying the etiology of childhood psychopathology using twin pairs the implications of possible contrast effects needed to be considered. As explained in the introduction, very low DZ correlations compared to MZ correlations and differences in variance for MZ and DZ twins give an indication that contrast effects are present. Both the variances and pattern of twin correlations for Internalizing and Externalizing behavior at ages 3, 7, 10, and 12 showed no indication of a contrast effect (see Chapter 3). Thus, contrast effects were not considered to be important for the results of the longitudinal analyses on the development of Internalizing and Externalizing problem behavior.

Rater bias and parental disagreement

As explained in the introduction, sources of rater bias are stereotyping, employing particular normative standards, of having a certain response style. Based on a specific rating tendency, parents will rate their children more alike than the really are. So the presence of rater bias will cause shared environmental effects to be overestimated. The analyses in Chapter 3, on the development pattern of genetic and environmental influences on Internalizing and Externalizing problem behavior were conducted with data from mother ratings only, so it could be the case that parts of the significant shared environmental influences on Internalizing and Externalizing behavior at a certain age and on the stability in both behaviors were partly based on rater bias. In order to distinguish 'real' shared environmental influences on problem behavior from rater bias, data based on mother and father ratings were used.

Two cross-sectional analyses have been conducted with mother and father ratings for Internalizing and Externalizing behavior in large samples of Dutch 10-and 12-year-old twins. The parental intercorrelation for both Internalizing and Externalizing behavior at ages 10 and 12 was around .60, which is in line with the results from a meta-analysis conducted by Achenbach and colleagues (1987). This high intercorrelation implied that both parents similarly assess part of the behavior. The intercorrelation, though, was less than perfect, indicating disagreement between parents. Differences in types of rater bias between raters leads to disagreement between raters. Another important source of disagreement is unreliability, which rises when raters cannot give an accurate description about relevant behaviors. Finally, it could be the case that parents do not assess exactly the same behavior in their children. It is know that different raters can provide, each from their own perspective, somewhat different but valid and complementary information about the child's functioning.

To disentangle sources of parental disagreement and to distinguish 'real' shared environmental influences from influences due to rater bias, distinct multiple rater models (introduced in Chapter 2) were fit to parental ratings of Internalizing and Externalizing behavior. It was found that rater differences do not merely reflect measurement error or rater bias, but indicate that parents assess different aspects of the child's behavior, as represented by the so-called Psychometric model. These results in 10-and 12-year-old twins are in accordance with previous studies (Hewitt *et al.*, 1992; Van der Valk *et al.*, 2001; Van der Valk *et al.*, in press). Neither for the Internalizing scale nor for the Externalizing scale measurement errors and unreliability account for more than 11% of the variance. Rater bias accounts for at most 13% of the variance for both the Internalizing and Externalizing scale.

An important finding from these cross-sectional studies was the significance of an unique view of each parent. These results can be linked to results from previous comparable studies in the 3-and 7-year old twins (Van der Valk *et al.*, 2001; Van der Valk *et al.*, in press). For Internalizing behavior, a possible specialization of the parent-child relationship over the years is represented by a relative increase of unique additive genetic factors, representing the parental unique view, from age 3 to age 12. At age 3 the unique additive genetic factors represented 16 % of the total additive genetic effects, while at age 10 the unique additive genetic effect explained 28% of total additive genetic variance based on mother ratings and 21% of the total additive genetic variance based on father ratings for Internalizing problem behavior. At age 12, for Internalizing behavior 41% of the total additive genetic variance was explained by unique additive genetic effects. When children grow older the mother-child and father-child relation may become more distinct, because of the fact that the child's behavior becomes more diverse over the years. The diversity of behavior may create more situational specific behavior, different for mothers and fathers.

For Externalizing behavior less change was observed. The relative importance of the unique view of the mother on the child's' behavior, was relatively stable over the years (around 15% in boys and around 25% in girls). Father's unique view, however, showed a remarkable drop at age 10 and 12. Fathers seem to add no unique view on Externalizing behavior at age 10 and 12 in both boys and girls. Future studies in this sample may be helpful to determine whether this effect will persist at older ages. These findings are of specific interest to the study of developmental psychopathology, because as children enter, endure, and exit puberty it will be important for parents to realize that they see and respond to different aspects of behavior. It is in fact the case, that the same child may appear different to mom and dad.

A final step in disentangling genetic and environmental influences on Internalizing and Externalizing problem behavior and in investigating the distinct developmental processes of these variance components was to study the developmental patterns of the parental shared and unique views. Further insight into the stability or change in the influences of rater bias and the parental unique view is essential in studying the developing child by making use of questionnaire data.

The longitudinal psychometric model

So far it has been established that genetic and shared environmental influences are important in explaining stability in Internalizing and Externalizing problem behavior throughout development. A small but significant part of the shared environmental influences on problem behavior seems to be accounted for by rater bias. Further,

disagreement between parents is result of rater bias, but parents also provide unique information on the child's behavior. The significance of the parental unique view was sorted out with the use of the Psychometric model. If the behaviors uniquely rated by the parents are shown to be influenced by the genotype of the child, the parent must have been assessing 'real' unique behavioral views. For error and/or unreliability cannot cause the systematic effects necessary for the model to estimate genetic influences.

To analyze longitudinal data on problem behavior as assessed by multiple raters a longitudinal Psychometric model was developed. This model provided the tools to simultaneously investigate the etiology of developmental patterns and plausibility of different models for (dis)agreement between multiple raters.

For Internalizing behavior a decrease in additive genetic influences on the reliable trait variance, representing behavior similar assessed by both parents, was found over the years, from 48% at age 3 to 28% at age 12. A complementary increase in common shared environmental influences from 6 % at age 3 to 20% at age 12 was found. Stability in Internalizing behavior, based on mother ratings was for 41% accounted for by common additive genetic factors and for 28% by common shared environmental influences. In Externalizing behavior, based on mother ratings, stability was for 53% accounted for by common additive genetic factors, while 23% was accounted for by common shared environmental influences. The developmental process for the common additive genetic, shared environmental and nonshared environmental influences were best described by a simplex structure.

An important finding is that the maternal and paternal unique shared environmental influences, partly, representing rater bias, were best described by a factor structure. This developmental structure indicates a significant influence of rater bias on stability in problem behaviors. In interpreting results of longitudinal studies it should be taken into account that part of the stability in the trait under investigation can be caused by stability in aspects outside the trait. In this case part of the stability is caused by stability in the 'rater bias' instead of stability in the 'real' behavior. Another important result of the analyses is the fact that the unique view of the father, represented by unique additive genetic influences, is time specific only.

Overall conclusions in studying the development of Internalizing and Externalizing problem behavior

Some overall conclusions can be drawn from the distinct but overlapping studies on Internalizing and Externalizing problem behavior in this project. Individual difference in Internalizing and Externalizing behavior at the distinct ages can be explained by additive genetic and shared environmental factors. For Internalizing behavior a decrease of genetic influences from 57% at age 3 to 22% at age 12 years was observed A complementary

increase in common shared environmental influences, from no significant influence at age 3 to 25% at age 12 years, was found. For Externalizing behavior the influences of additive genetic and shared environmental factors remained relatively stable over the years explaining about 45% and 20% of the total variance, respectively.

The significant influences of additive genetic factors on the reliable trait variance indicate a possible innate vulnerability to childhood psychopathology. The influences of nonshared environmental factors, explaining 10 to 20% of the total variance in both Internalizing and Externalizing behavior suggest the importance of pure idiosyncratic experiences. Significant influences of environment shared by both members of a twin pair, like home environment, are represented by the common shared environmental factor.

Table 12.3.

Decomposition of the total variance and covariances for Internalizing behavior based on a single and a multiple rater model

		Variance				Covariance					
		Int3	**Int7**	**Int10**	**Int12**	**Int3-Int7**	**Int3-Int10**	**Int3-Int12**	**Int7-Int10**	**Int7-Int12**	**Int10-nt12**
Sinlge Rater	**A**[a]	.59	.44	.36	.37	.52	.43	.32	.51	.40	.38
	C[b]	.13	.28	.34	.37	.43	.56	.63	.36	.42	.40
	E[c]	.28	.28	.31	.26	.05	.01	.04	.13	.17	.22
Multiple Raters	**A**[d]	.48	.36	.28	.28	.50	.45	.45	.39	.34	.31
	A$_m$[e]	.13	.12	.11	.11	.02	.02	.00	.09	.04	.09
	C[f]	.06	.13	.19	.20	.24	.29	.35	.22	.29	.27
	C$_m$[g]	.05	.15	.14	.17	.19	.24	.24	.17	.16	.12
	E[h]	.14	.15	.15	.14	.03	.01	.00	.11	.12	.16
	E$_m$[i]	.14	.12	.15	.12	.02	.00	.05	.03	.06	.06
	A[d]	.49	.35	.30	.27	.59	.56	.54	.47	.39	.36
	A$_f$[i]	.09	.12	.07	.00	.08	.09	.00	.00	.00	.02
	C[f]	.06	.14	.21	.22	.28	.36	.42	.26	.33	.31
	C$_f$[k]	.09	.12	.15	.08	.02	.06	.18	.15	.15	.13
	E[h]	.14	.16	.17	.15	.04	.00	.00	.14	.13	.18
	E$_f$[l]	.12	.11	.10	.11	.00	.00	.01	.01	.02	.00

The importance of the use of multiple raters to reliable estimate the influences of genetic and environmental factors on variance and covariances of problem behavior can be seen in Table 12.3 (Internalizing behavior) and 12.4 (Externalizing behavior).

Comparing estimates of genetic and environmental influences on covariance using single rater models and multiple rater models showed that a significant fraction of the shared environmental influences on stability were accounted for by parental unique shared environmental influences (C_m and C_f), which partly represent rater bias. So by interpreting results of longitudinal studies it should be noted that the stability is partly accounted for by stability in rater instead of stability in 'real' behavior. The unique additive genetic influences (A_m and A_f) represent the finding of a unique view of each parent on his or her child's behavior. This result implies that each parent provides additional information on the child's functioning and in that matter the use of multiple rater to study individual differences at distinct ages is recommended. However, when studying developmental patterns to predict and prevent maladjustment the use of father ratings is questionable. Results of this study indicated that the influences of the paternal specific view (Af) were time specific only. Fathers seem to attribute additional information on the child's behavior, however no continuity in this view was observed. It could be the case that these age specific paternal views are a representation of trivial behavioral fluctuations rather than real father-child specific interaction. Further, mothers' specific view showed continuity over the years, the influence on the covariance, though, was very low. In the scope of aim of the study and actual resources of time, money and practical attainability it could be reasonable to use mother ratings of genetically related individuals only in studying the influences of genetic and environmental factors on the development of problem behavior. However, by doing so, one should always be aware of the significant influences of rater bias, resulting in an overestimation of shared environmental influences on the development of behavioral problems.

Finally an important feature of the present longitudinal twin studies was the possibility to investigate the developmental pattern of genetic and environmental factors separately. Influences of additive genetic, shared and nonshared environment on the development of problem behavior were best described by a simplex pattern. This simplex-like continuity for genetic and environmental influences assumes that successive levels of functioning were causally linked and that earlier experiences and/or genetic effects affected later maladjustment. Nonshared environmental influences seemed to be mainly important for age-specific behavior problems and have almost no developmental significance. This implies that influences of nonshared environment are important but that they are mostly of transient nature and specific to a specific moment in time.

Table 12.4.

Decomposition of the total variance and covariances for Externalizing behavior based on a single and a multiple rater model

		Variance				Covariance					
		Int3	Int7	Int10	Int12	Int3-Int7	Int3-Int10	Int3-Int12	Int7-Int10	Int7-Int12	Int10-nt12
Sinlge Rater	A^a ♂	.57	.59	.65	.64	.64	.66	.57	.74	.69	.73
	♀	.50	.59	.45	.51	.54	.45	.44	.60	.58	.58
	C^b ♂	.27	.26	.20	.23	.33	.33	.38	.19	.23	.18
	♀	.32	.27	.36	.33	.38	.50	.52	.32	.33	.33
	E^c ♂	.16	.15	.14	.13	.03	.01	.05	.07	.07	.08
	♀	.18	.14	.19	.16	.07	.06	.04	.09	.09	.09
Multiple Raters	A^d ♂	.48	.49	.52	.47	.58	.55	.51	.55	.58	.61
	♀	.41	.46	.41	.40	.48	.41	.41	.48	.52	.52
	A_m^e ♂	.05	.11	.12	.13	.04	.06	.07	.12	.13	.08
	♀	.11	.13	.09	.11	.08	.13	.08	.09	.06	.11
	C^f ♂	.18	.15	.14	.19	.20	.25	.25	.20	.17	.16
	♀	.22	.17	.19	.21	.27	.34	.34	.25	.24	.21
	C_m^g ♂	.13	.09	.07	.09	.14	.13	.13	.04	.06	.06
	♀	.08	.10	.13	.13	.10	.08	.13	.08	.09	.08
	E^h ♂	.10	.09	.08	.06	.04	.02	.02	.06	.05	.06
	♀	.11	.09	.10	.09	.05	.05	.03	.09	.06	.08
	E_m^i ♂	.07	.06	.06	.06	.00	.00	.02	.01	.02	.03
	♀	.07	.06	.08	.07	.02	.00	.00	.01	.03	.02
	A^d ♂	.48	.50	.53	.49	.66	.60	.59	.61	.67	.66
	♀	.42	.48	.43	.43	.57	.45	.50	.51	.58	.57
	A_f^j ♂	.05	.09	.02	.02	.03	.00	.00	.03	.02	.00
	♀	.04	.03	.05	.02	.00	.00	.00	.05	.00	.00
	C^f ♂	.18	.16	.15	.20	.23	.27	.29	.22	.19	.17
	♀	.23	.18	.20	.22	.31	.38	.41	.27	.27	.23
	C_f^k ♂	.11	.11	.16	.14	.03	.13	.10	.06	.06	.14
	♀	.13	.15	.14	.15	.08	.14	.08	.08	.10	.12
	E^h ♂	.10	.09	.09	.07	.04	.02	.02	.07	.06	.07
	♀	.11	.09	.11	.09	.06	.06	.04	.10	.07	.08
	E_f^l ♂	.08	.06	.05	.04	.02	.00	.03	.00	.05	.00
	♀	.07	.07	.07	.08	.00	.00	.00	.00	.00	.03

[a] Additive genetic influences; [b] Shared environmental influences; [c] nonshared environmental influences; [d] Additive genetic influences on the reliable trait variance; [e] Maternal unique Additive genetic influences; [f] Shared environmental influences on the reliable trait variance; [g] Maternal unique shared environmental influences; [h] nonshared environmental influences on the reliable trait variance; [I] Maternal unique nonshared environmental influences; [j] Paternal unique additive genetic influences; [k] Paternal unique shared environmental influences; [l] Paternal unique nonshared environmental influences

Cognitive abilities

Heritability of cognition has been studied extensively, both in adults and in children. Far less is known about the developmental genetics of cognitive abilities. In this project the influences of genes and environment on cognitive development and on its developmental structure were studied in a longitudinal sample of Dutch twins at 5, 7, 10, and 12 years of age. It can be concluded that the development of general cognitive abilities is a continuous process, represented by high correlations over time ($r_{(5-7)}$= .65; $r_{(5-10)}$= .65; $r_{(5-12)}$=.64; $r_{(7-10)}$=.72; $r_{(7-12)}$=.69 and $r_{(10-12)}$=.78). Continuity was represented by a common factor, with age specific factor loadings, for both genetic and shared environmental influences. Change in development, represented by age specific factors, was presented in the shared environmental structure and, as expected, in the nonshared environmental structure. Further, decomposition of the between age covariances in additive genetic, shared environmental, and nonshared environmental influences showed that the continuity in cognitive abilities was mainly due to additive genetic factors, accounting for 72% of the covariance on average. The remaining 28% could be explained by environmental influences shared by both members of a twin pair.

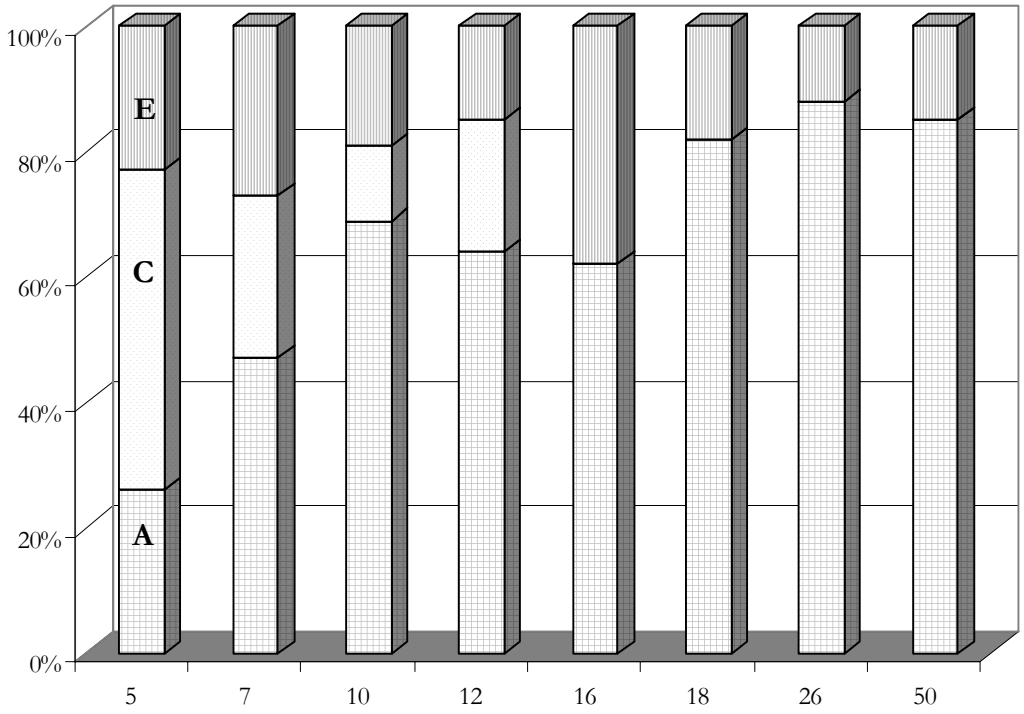
An increase in heritability of IQ over the years was found. At age 5, 26% of the total variance was accounted for by genetic influences, while at age 10, 69% of the total variance was explained by genetic factors. Estimates on heritability and influences of shared and nonshared environment based on several studies with twin from the NTR (Rietveld *et al.*, 2002; Posthuma, 2002; Van Baal, 1997; Rijsdijk; 1997a; Van Beijsterveldt, 1996) are summarized in Diagram 12.1. The increase of additive genetic influences is clearly observed, so is the decrease of shared environmental influences.

Based on previous longitudinal studies (Cardon *et al.*, 1992; Fulker, *et al.*, 1993) the finding of a common factor for shared environmental influences was expected. However, beside this common factor, age specific influences of shared environment were found. These age specific influences were significant, but the proportion of variance explained by this shared environmental factors was much smaller compared to the proportion explained by the shared environmental factor common to all ages. SES and parental education could account for this common factor, as these environmental aspects are not sensitive to large changes over a time-span of 7 years. For the age specific shared environmental influences one may consider the school environment. In 63% of a large sample of Dutch 12-year old twins, the same teacher teaches both children of a twin pair, whereas 37% go to separate classes. This ratio makes teacher or classroom environment a shared environmental influence for the majority of the children. Since, in the Dutch school system children move to a different teacher each school year, this results in a lack of continuity in this particular

aspect of shared environment. So, these shared but age-specific experiences within the classroom may be represented by the age-specific factors.

Diagram 12.1.

Influences of additive genetic (A), shared environment (C), and nonshared environment (E) on cognitive abilities at distinct ages (Data at ages 16 to 50 are provided by the Netherlands Twin Register)



The nonshared environment was found to explain a substantial portion of the variance at each age (best model, range from 15% to 27%). With respect to developmental aspects of the data the nonshared environment acts in a well-established manner. The environment that is uniquely experienced by an individual contributes to change rather than stability in cognitive performance.

The developmental pattern for genetic influences found in this study is partly different from previous, comparable studies like the combined study of CAP, MLTS, and TIP (Cardon *et al.*, 1992; Fulker, *et al.*, 1993). Results provided by these studies show a simplex pattern for genetic influences with genetic innovation at 2, 3, and 7 years of age. In our study no indication for genetic innovation is observed. Our result point into the direction of a common set of genes influencing cognitive abilities throughout development. It can be

hypothesized that differences in results from distinct studies can be partly explained by differences in age of the subjects and methods to assess cognitive abilities.

Educational achievement and the overlap with cognitive abilities

Intelligence seems to be an obvious explaining factor for the variance in educational achievement. In this project significant correlations were found between IQ and a national test of educational achievement (CITO). The correlations were .41, .50, .60, and .63 between CITO and IQ assessed at ages 5, 7, 10, and 12 years, respectively. Additive genetic effects accounted for 60% of the individual differences found in CITO scores. This high heritability indicates that the CITO might be a valuable instrument to assess individual differences in cognitive abilities in children but might not be the right instrument to put the effect of education to the test. The results of the bivariate analyses pointed to genetic effects as an important source of covariance between CITO and IQ, explaining 60% of the covariance on average. Shared environmental influences seemed significant in explaining the overlap between scholastic achievement and IQ as well.

In general, the family environment (SES) is considered to be the main factor of shared environmental influences. As mentioned in the study on the development of cognitive abilities, one may also consider the school environment as an important source of shared environmental influences, which can be age specific due to the move to a different teacher each school year. Further indication to consider the classroom and teacher as shared environment is given by preliminary results on twin correlations for CITO in the same sample of 12- year- old-twin pairs. The pattern of twin correlation for CITO in twins taught by the same teacher indicated higher influences of shared environment than the pattern of twin correlations for CITO in twins taught by different teachers. Two limitations of the preliminary results must be emphasized. First, a problem is that there may be reasons why twins go to separate classes. For instance, educational achievement of both children of a twin pair diverges over the years, resulting in different classes. Based on the high heritability in both cognitive abilities as well as scholastic achievement this is more likely to occur in DZ twins. Further, it should be noted that because only a minority of the twin go to separate classes the zygosity groups to calculate these twin correlation are very small. The collection of CITO data and data on 'different or same' teacher is a continuous process at the NTR, so more insight into this matter can be gained in the future.

The unique environment was found to explain only a small portion of the variance of CITO. It further seems to be of no influence on the association between CITO and IQ, except for the association between CITO and IQ12. The finding of the influence of nonshared environmental influences on this covariance indicates that, besides measurement error, pure idiosyncratic experience are of importance for individual

differences in cognitive abilities and CITO at age 12. Further, the finding of a significant influence on the association between CITO and IQ at age 12 only, underlines the transient nature of these idiosyncratic experiences. This transient nature of nonshared environmental influences was also found in the longitudinal study on the development of intelligence.

The finding of genetic influences as the overlapping factor for the association creates opportunities for future research on the genetics of cognition. Assessing an intelligence test is very time consuming and in order to get more insight in the genetic background of cognition large sample sizes are necessary. Since the CITO is a nationwide standardized test, the use of the database and the possibilities to recruit parents, siblings and normal controls for genetic studies would boost power to finally find genes influencing cognitive abilities.

Cortisol

The recent literature on the heritability of cortisol levels was reviewed and it was observed that most of the studies, which have been carried out in genetically informative samples, lack methodological consistency with regard to frequency and timing of sample collection. The circadian rhythm in cortisol levels was often not taken into account. A power analysis showed that none of the reviewed studies used adequate sample sizes to distinguish genetic from shared environmental influences as a cause for familial aggregation. Results of a simultaneous analysis of 5 comparable twin studies suggested a heritability of 62% for basal cortisol levels. Hence, it was concluded that, to understand the contribution of genetic and environmental influences to variation in basal cortisol levels, future studies should be designed more rigorously with strict collection and sampling protocols, sufficient sample size and repeated measures across multiple days.

Taking these recommendations into account a study on the heritability of cortisol in 12-year-old children was conducted. To this end, four samples of salivary cortisol were collected on two consecutive days in a sample of 180 twin pairs. The results showed a significant genetic contribution to the variation of basal cortisol levels in the morning and afternoon samples. Heritability did not differ for boys and girls and was highest (around 60%) for cortisol levels during the morning.

Rapid progress in understanding the genetics of cortisol receptors is made through pharmacological and knockout studies in animals. However, although large homology probably exists between animals and human HPA genes, genetic linkage or candidate gene studies in humans may well be needed. The heritability of cortisol levels found here suggests that this may be a useful phenotype for future gene hunting projects. Nevertheless, before cortisol can be adopted as an endophenotype for molecular genetic

studies, it is clearly essential to demonstrate that this measure of cortisol levels is strongly related to psychopathology or cognition, that they are influenced genetically, and that genes that influence cortisol also are likely to be involved in the etiology of behavioral problems or cognitive abilities. The finding of low and insignificant correlations between basal cortisol levels and problem behavior or cognition questions to what extent basal cortisol levels can be considered the designated endophenotype to explain individual differences in psychopathology and cognition.

FUTURE RESEARCH

The results of this study in light of future gene finding

The elusive goal of behavior geneticists will be the isolation of genes mediating complex behavioral phenotypes, such as psychopathology or cognition. It is established that behavior is partly inherited, by transmission of alleles from one generation to the next. The prevalence and magnitude of environmental influences, tough, tend to obscure behavior differences between distinct genotypes. A first step in successfully searching for genes influencing complex behavior (QTL; quantitative trait loci) is to disentangle these genetic and environmental influences. A next step will be the determination of the number and nature of inherited factors contributing to a behavioral trait and the mapping of genetic factors to positions on chromosomes. The final step will be the determination of how genetic factors function to generate behavior. Given the rapid advancements made in molecular biology, the development of statistical genetic methods, and the sequencing of the human genome, the identification of specific genes, even for complex traits, becomes a realistic goal of quantitative genetic analyses.

Despite these rapid advancements, though, a major obstacle to identifying genes for cognition and psychopathology remains the issue of phenotype definition. In this matter longitudinal studies in genetically related individuals, like this study, are valuable, because these studies take into account the magnitude and developmental pattern of genetic and environmental factors, the age of the subject, the gender of the subject, an the informant used. Further, although the heritabilities in cognitive abilities and psychopathology are high, the genetic influences are likely to be determined by a complex interplay of multiple biological and physiological processes, each influenced by its own set of genes. So, one approach to identify genes for complex traits could be the use of endophenotypes. In general endophenotypes represent biological, neurophysiological, electrophysiological, and behavioral indices of the pathways that connect genes and the trait under investigation. Based on several hypothesis (see Chapter 11) cortisol could be considered as an endophenotype for individual differences in psychopathology and cognition. In this project more insight into the etiology of individual differences in cortisol levels is gained. Future

studies, though, are necessary to estimate the value of the use of basal cortisol as an endophenotype to identify genes for psychopathology and cognition.

The development pattern of additive genetic influences on cognition was best captured by a common factor structure. This finding of a common set of genes has important implications for the search for genes for cognition. If the same genes influence cognition at distinct ages, the search for QTLs can be conducted in a group of subjects of varying ages. For Internalizing and Externalizing behavior, though, additive genetic influences on development, were best described by a simplex pattern. This structure implies that genetic variance is partly transmitted from one age to the next, but new genetic influences come into play at each age as well. Thus, the set of genes influencing Internalizing and Externalizing behavior is not common to all ages and DNA of distinct age groups need to be considered for finding QTLs.

Finally, an important approach for investigating genetic influences on complex traits involves examining the consequences of experimental manipulations of genes. Because this work cannot be performed in humans, model organisms much be used for such studies. Genetic regulation of behavior may be explored in the mouse by examining naturally occurring gene variants and by introducing genetic mutations. Moreover, humans and mice possess remarkably similar genomes (Makalowski and Boguski, 1998*)*. Therefore, mouse genetic studies can provide insight into the actions of corresponding human genes. Collaboration of human and animal studies might facilitate and accelerate the finding of genes for complex behaviors.

The genetic basis of psychopathology

Most of the studies on the genetics of childhood psychopathology focus on attention deficit hyperactivity disorder (ADHD) and aggression. This attention may be partly due to the high heritability found for these disorders. In this project focus has been on Internalizing and Externalizing behavior as assessed with the Child Behavior Checklist. Syndromes used to define Externalizing behavior are Attention problems and Aggressive behavior. Externalizing behavior is a valuable concept to gain insight into development of behavioral problems. For the finding of genes influencing these kinds of behavior, though, this broadband scale may be to complex.

Replicated findings for susceptibility genes for ADHD are reported for the dopamine (DRD4) receptor gene and the dopamine transporter gene (DAT1). The indication of the involvement of the dopaminergic system in individual differences in ADHD is based on the fact that 70 to 80% of children with ADHD show an immediate improvement in ADHD symptoms when given stimulant medication such as methylphenidate (e.g. Ritalin). These genes are known to inhibit reuptake via the dopamine transporter (Amara and

Kuhar, 1993) and increase synaptic levels of dopamine (Solanto, 1998). Several studies have examined the association of a genetic polymorphism (the 48bp VNTR in exon 3) of the DRD4 gene with ADHD. Eight independent studies have found evidence of association of DRD4 7-repeat allele with ADHD (Lahoste *et al.*, 1996; Smalley *et al.*, 1998; Swanson *et al.*, 1998; Comings *et al.*, 1999; Faraone *et al.*, 1999; Muglia *et al.*, 2000; Sunohara *et al.*, 2000; Tahir *et al.*, 2000). A recent meta-analysis demonstrated significant association (and linkage) of the DRD4 repeat allele with ADHD from both family-based studies and case control studies (Faraone *et al.*, 2001) . Thus overall, so far evidence suggests that there is association between the DRD4 7-repeat allele and ADHD. Further, there have been at least eight published studies of DAT1 and ADHD, all of which have examined the same genetic variant, a VNTR at the 3' region. Four studies have shown significant linkage and association with allele 10, the 480 bp repeat (Cook *et al.*, 1995; ; Gill *et al.*, 1997; Daly *et al.*, 1999; Curran *et al.*, 2001). One study found evidence of a trend for association (Barr *et al.*, 2001). Three studies, though, failed to show association (Swanson *et al.*, 1998; Palmer *et al.*, 1999; Holmes *et al.*, 2000).

Besides the interest for the dopamine hypothesis, more recently the potential role of seretonin (5-HT) in the etiology of ADHD has been highlighted. This has been suggested by findings from animal studies and the effect that stimulant medication and second-line therapeutic drugs for ADHD have on noadrenergic pathways (Solanto, 1998). In general, 5-HT appears to play a role in a range of neuropsychiatric disorders. For instance, genes encoding various components of the 5-HT system are being studies as risk factors in depression, obsessive-compulsive disorder, and aggression (for a review see Lucki, 1998).

Further, genetic deficiencies in *MAOA* have been linked with aggression in mice and humans (Rowe, 2001; Manuck *et al.*, 2000). The *MAOA* gene is located on the X chromosome (Xp 11.23-11.4) and it encodes the *MAOA* enzyme, which metabolizes neurotransmitter such as norepinephrine (NE), seretonin (5-HT), and dopamine (DA). Increased aggression and increased levels of brain NE, 5-HT, and DA were observed in a transgenic mouse line in which the gene encoding *MAOA* was deleted (Cases, 1995), and aggression was normalized by restoring *MAOA* expression (Shih and Thompson, 1999). In humans, a null allele at the *MAOA* locus was linked with male antisocial behavior in a single large family studied in the Netherlands (Brunner *et al.*, 1993). However, this mutation is very rare. Further preliminary evidence of an association between polymorphic variation in the gene for *MAOA* and interindividual variability on aggression, impulsivity and central nervous system serotonergic responsitivity is found by Manuck and colleagues (2000). Recently, a significant G×E interaction was reported between *MAOA* and maltreatment. Caspi and colleagues (2002) found that maltreated children with a genotype conferring high levels of *MAOA* expression were less likely to develop antisocial behavior. Overall,

evidence for an association between *MAOA* and aggressive behavior or antisocial behavior in human general population remains inconclusive and replication of the results are needed.

The genetic basis of cognition

In spite of the overwhelming evidence for the moderate to high heritability in cognitive abilities, actual genes have not yet been identified. A recent review of the exiting literature presents a list of over 150 candidate genes that contribute to cognition (Morley and Montgomery, 2001). The tumor-suppressor gene NF1, which encodes a Ras-specific GTPase activating protein, has been implicated in the cognitive processes of humans, mice and *Drosophilae*. In humans this gene is linked to neurofibromatosis type 1 (Hofman *et al.*, 1994; North *et al.*, 1995; Ferner *et al.*, 1996; Silva *et al.*, 1997; Ozonoff *et al.*, 1999), a disease that includes a number of general cognitive impairments, and Watson syndrome, an autosomal dominant condition that includes low intelligence (Allanson *et al.*, 1991). The ADRA2C (adrenergic alpha 2A receptor) gene is a strong candidate for a role in human cognition due to its involvement in attention, learning and memory. The additive effects of ADRA2C and dopamine beta-hydroxylase genes were significantly associated with ADHD, learning disabilities and poor scholastic performance in children (Comings *et al.* 1999). Finally, there is evidence for a role of APOE in human and mouse learning and memory. The presence of at least one epsilon 4 allele is associated with lower IQ scores, especially with lower verbal IQ (Posthuma *et al.*, 1999a; 1999b). Henderson and colleagues (1995) showed that the epsilon 4 allele of APOE is associated with increased risk for dementia or cognitive impairment in aged subjects. Further, subjects with APOE epsilon 4 allele have decreased learning and memory abilities compared with control subjects (Schmidt *et al.*, 1996).

Finding genes associated with complex quantitative traits, such as cognition, requires power to detect QTLs of small effect size. Recently, a genome wide scan for cognitive abilities was conducted, using 1842 markers across the genome (Plomin *et al.,* 2001). In order to detect QTLs of small effects size, they used extreme selected samples and five-stage design with normal alpha levels that permit false positive results in early stages but remove false positives in later stages. Despite this approach they could not replicate any of the previously found QTL associations and did not detect new QTL associations.

To summarize

In studying individual differences in development during childhood insight is gained into the etiology of human variation. Behavioral problems and cognitive abilities are stable over the years and genetic and shared environmental influences are important in explaining this

stability. A comprehensive picture of developmental processes underlying child psychopathology and cognition is obtained. Important to reliable estimate the influences of genetic and environmental factors on problem behavior is the use of information from multiple raters. Further, It was hypothesized that individual differences in pre- or postnatal cortisol levels could be an indirect cause of individual differences in psychopathology or cognition. The finding of low and insignificant correlations between basal cortisol levels and problem behavior or cognition questions the value of this hypothesis.

Summary

This thesis describes the outcome of two longitudinal studies. Two important phenotypes are considered in a large longitudinal sample of Dutch twins: cognitive abilities and childhood psychopathology. Cognitive abilities were studied in a longitudinal sample of 400 children. Measures of intelligence were collected at 5, 7, 10, and 12 years of age. Additional information on cognitive abilities at age 12 was collected by means of the CITO score for 1495 children. Behavioral and emotional problems were assessed longitudinally, by parental report, in over 10,000 children at ages 3 and 7, in 6000 children at age 10 and in 3000 children at age 12. Finally, salivary cortisol samples were collected at two consecutive days in 180 twelve-year-old twin pairs, who participated in the longitudinal study on cognitive abilities.

The longitudinal study on cognitive abilities

It can be concluded that the development of general cognitive abilities is a continuous process, represented by high correlations over time ($r_{(5-7)}$= .65; $r_{(5-10)}$= .65; $r_{(5-12)}$=.64; $r_{(7-10)}$=.72; $r_{(7-12)}$=.69 and $r_{(10-12)}$=.78). Stability in cognitive ability throughout development is mainly accounted for by genetic factors. Seventy-two percent of the between age covariance, representing stability in cognitive abilities over time, was accounted for by additive genetic factors. The remaining 28% could be explained by environmental influences shared by both members of a twin pair. The developmental pattern of additive genetic influences (A) on cognition is best described by a common factor structure. Thus, an underlying set of genes influences cognition from the beginning onwards. However, the importance of these genes increases greatly during childhood (see Figure I). Environmental influences shared by both members of a twin pair, the so-called shared environmental influences (C), are less important over time. The developmental pattern is also best captured by a common factor structure. Besides this common factor structure, significant age specific influences of shared environment in cognitive abilities are found. Nonshared environmental influences, environmental influences unique for each person (E), show no continuity, thus mainly account for change in cognitive functioning over the years.

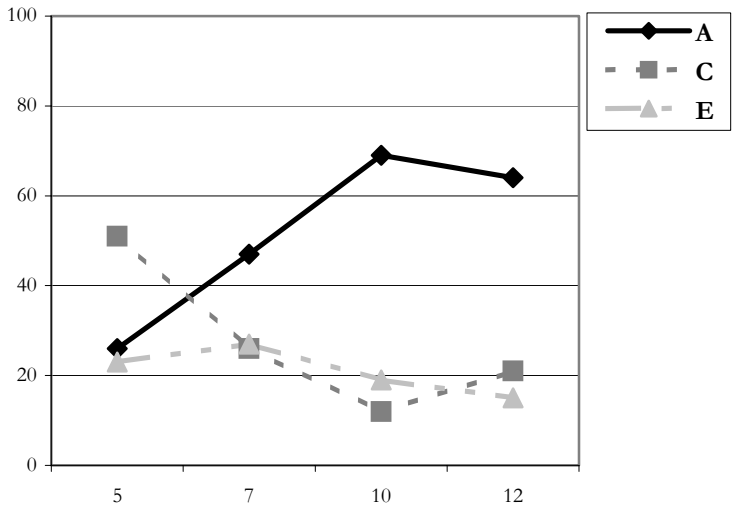
As depicted in figure I, an increase in heritability of IQ over the years was found. At age 5, 26% of the total variance was accounted for by genetic influences, while at age 10, 69% of the total variance was explained by genetic factors. A complementary decrease in shared environmental influences is observed.

The correlation between scholastic achievement, assessed with the CITO score at age 12, and cognitive abilities at ages 5, 7, 10 and 12 years (.41, .50, .60, and .63 resp.) is explained by genetic and shared environmental factors. Individual differences in CITO score are mainly accounted for by additive genetic influences, explaining 60% of the total

variance. Small but significant influences of shared environment are found, explaining about 25% of the total variance. The remaining 15% is accounted for by nonshared environmental influences.

Figure I.

Graphical representation of the variance decomposition in additive genetic (A), shared environmental (C), and nonshared environmental (E) factors for cognitive abilities at age 5, 7, 10, and 12 years.



The longitudinal study on behavioral and emotional problems

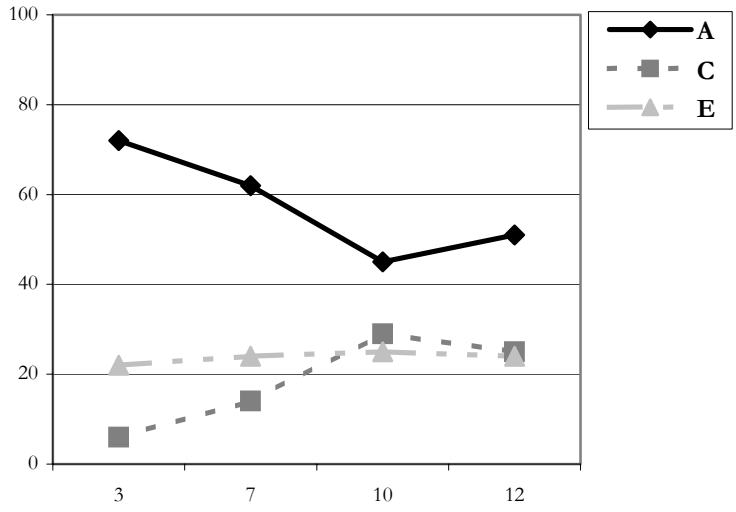
Behavioral and emotional problems are highly prevalent in Dutch children aged 3 to 12 years. Twenty-five to 30% of the children in the sample showed some kind of behavioral problem. The prevalence of behavioral problems, though, varies depending on the kind of behavior, the rater, and the age and gender of the child. Internalizing problems (anxious/depressed behavior, withdrawn behavior) are more prevalent in girls than boys and Externalizing problems (aggressive behavior, rule breaking behavior) are more prevalent in boys than girls. Overall, mothers report more problem behaviors than fathers and prevalence of problem behavior, as rated by the parents, decreases over age.

As for cognitive abilities, it can be concluded that the development of Internalizing and Externalizing problem behavior is a continuous process, represented by moderate to high correlations over time. The observed stability coefficients for the four- seven-, and nine-year time intervals were, respectively, .37, .33, .30 for Internalizing behavior and .55, .49, and .48 for Externalizing behavior. Stability in problem behaviors can be explained by genetic and environmental influences on the covariances over time. For Internalizing

behavior in boys stability is for 65%, 26%, and 9% explained by additive genetic, shared environmental and nonshared environmental influences, respectively. Forty-seven percent of the stability in girls is accounted for by additive genetic factors. Stability in Internalizing behavior in girls is accounted for 43% by shared environmental factors. The remaining 10% can be explained by nonshared environmental factors. Genetic influences are the main factor for stability in Externalizing behavior in boy, explaining 76% of the covariance over time. Nineteen percent is accounted for by shared environmental influences. Only 5 % of the stability in Externalizing behavior in boys is accounted for by nonshared environmental factors. For Externalizing behavior in girls, both genetic and shared environmental factors are important for stability over the years, explaining 62% and 31% respectively. The remaining 7% is accounted for by nonshared environmental factors. For both Internalizing and Externalizing behavioral problems the developmental pattern of genetic, shared and nonshared environmental influences is best described by a simplex pattern. So, influences are transmitted from age to age and new influences come into play at each age.

Figure II.

Graphical representation of the variance decomposition in additive genetic (A), shared environmental (C), and nonshared environmental (E) factors for Internalizing behavior in boys, similar assessed by both parents, at age 3, 7, 10, and 12 years.

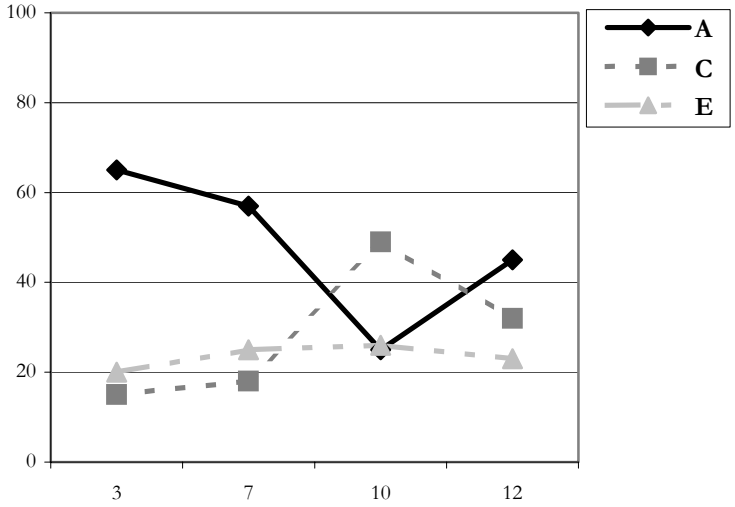


For Internalizing behavior (See Figure II) in both boys and girls a decrease in genetic influences and a complementary increase of shared environmental influences is observed.

Relative stability in the strength of genetic and environmental influences in boys is reached from age 10 onwards.

Figure II-cont.

Graphical representation of the variance decomposition in additive genetic (A), shared environmental (C), and nonshared environmental (E) factors for Internalizing behavior in girls, similar assessed by both parents, at age 3, 7, 10, and 12 years.



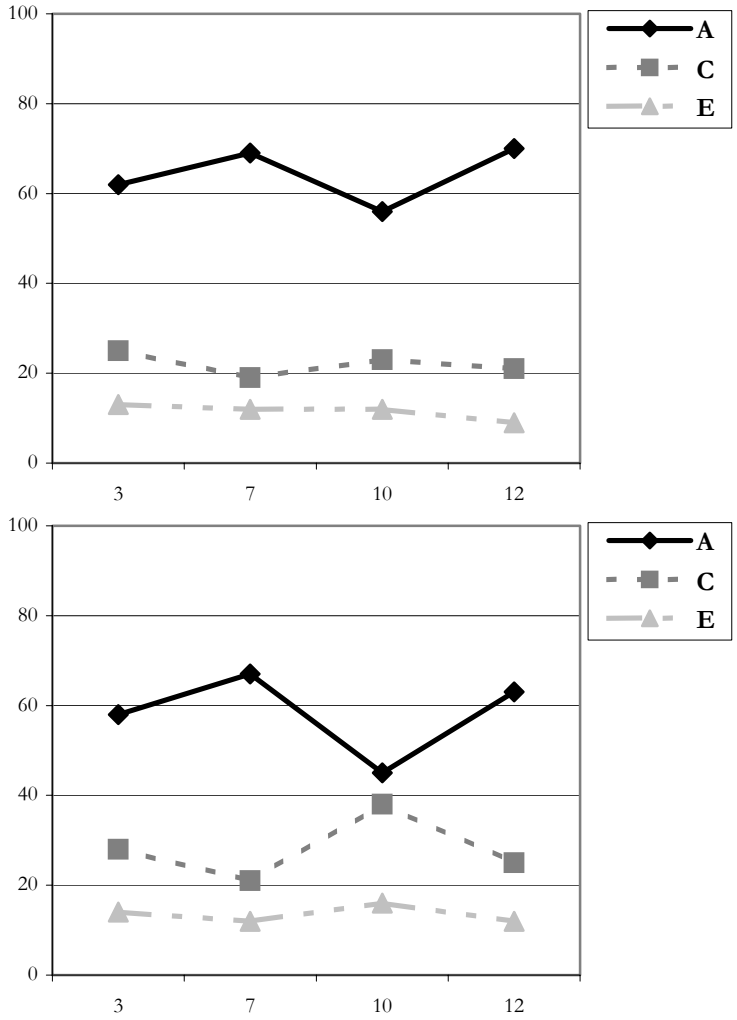
For Externalizing problem behavior less change in the strength of genetic and environmental influences is observed (see Figure III). Stability in Externalizing problem behaviors throughout development in boys and girls is mainly accounted for by additive genetic influences. For both Internalizing and Externalizing behavior it can be seen that the influences of nonshared environmental influences, representing pure idiosyncratic experiences, remains stable throughout development.

It can be concluded from this study that the developmental processes of genetic and environmental influences on the different phenotypes (cognition, Internalizing behavior, and Externalizing behavior) show distinct patterns. While a common set of genes influences cognition during childhood, a transmission process including new genetic influences at distinct ages best describes genetic influences on behavioral problems during childhood. Further, the strength of genetic and environmental influences changes during childhood, with an increase of genetic influences on cognition, a decrease of genetic

influences on Internalizing behavior and relatively stable genetic influences on Externalizing behavior.

Figure III.

Graphical representation of the additive genetic (A), shared environmental (C), and nonshared environmental (E) influences on Externalizing problem behavior, similar assessed by both parents, for boys (upper diagram) and girls (lower diagram) throughout development.



The use of multiple raters in studying the development of both problem behaviors is recommended. Agreement between parents for each assessment (age 3, 7, 10 and

12) was about .6, which indicates that part of the behavior is similar assessed by mothers and fathers. It is further found that disagreement between the parents, indicated by the less than perfect agreement, is not merely the result of unreliability and/or rater bias, but each parent also provides unique information from his/her own perspective on the child's behavior. These parental unique views show almost no continuity over the years, so this parent specific view is important for behavioral assessment at a certain age but is not significant for understanding the development of problem behavior during childhood. However, about 20% of the stability in Internalizing and Externalizing problem behavior over the years is accounted for by so-called rater bias. In other words, a part of the observed stability in problem behavior is accounted for by characteristics of the rater instead of factors of 'real' behavior. This finding emphasizes the importance of the use of multiple raters in studying the development of problem behavior.

Cortisol is mainly known for its pivotal role in generating an adequate response to physical and emotional stressors. However, it may also exert strong behavioral effects that are already apparent during childhood. In general biological, neurophysiological, electrophysiological, and behavioral indices of the pathways that connect genes and the trait under investigation, such as cortisol, are called endophenotypes. Before cortisol can be considered as an endophenotype in studying individual differences in the development of cognitive abilities and psychopathology, it is clearly essential to demonstrate that this measure of cortisol levels is strongly related to psychopathology or cognition, that both phenotypes are influenced genetically, and that genes that influence cortisol also are likely to be involved in the etiology of behavioral problems or cognitive abilities. Previous studies report an association between cortisol and cognitive abilities or childhood psychopathology. Further, the significance of genetic influences on childhood psychopathology and cognitive abilities are well established. However, more insight into the etiology of individual differences in cortisol levels, especially in children, still need to be gained.

For the determination of cortisol levels in children four saliva samples were collected at two consecutive days in a subsample (N=180 twin pairs) of the sample that participated in the longitudinal study of cognitive abilties. Results of the analyses showed a significant genetic contribution to the variation of basal cortisol levels in the morning and afternoon samples. Heritability did not differ for boys and girls and was highest (60%) for cortisol levels during the sample taken about 45 minutes after awakening. This cortisol awakening response provides a useful endophenotype in the search for genes that may affect hypothalamic-pituitary adrenocortical functioning in children.

Low correlations were found between cortisol levels at different points in time during the day. It was hypothesized that the heritability of cortisol levels varies inversely with the strength of the negative feedback signal exerted by cortisol at the glucocorticoid receptors (GR) and mineralocorticoid receptors (MR). Changes in the strength of this feedback signal are reflected in changes in the absolute cortisol level, although time lagged, because the effects of cortisol on the GR and MR receptors are largely genomic.

Preliminary results show no significant correlation between cortisol levels assessed at age 12 and Internalizing or Externalizing behavior at ages 3, 7, 10, and 12. Further, no significant association is observed between cognitive abilities at ages 5, 7, 10, and 12 and cortisol levels at age 12. Future studies are necessary to gain more insight into the presence or absence of these associations and the possible value of cortisol levels as an endophenotype in studying individual differences is the development of cognition and childhood psychopathology.

Based on the results of this thesis it can be concluded that genetic influences are pervasive during childhood. Genes affect two major phenotypes, behavioral problems and cognitive abilities, as well as a possible endophenotype, basal cortisol levels. Genes also affect various aspects of development and are very important for continuity and change throughout development. It can be stated that genetic studies are not only important for estimating size of genetic and environmental effects, but can shed light on fundamental questions in child development. For this reason, genetic studies supply phenotypic studies and are an essential addendum for research in developmental psychology.

Samenvatting

Dit proefschrift beschrijft de resultaten van twee longitudinale en een cross-sectioneel onderzoek. Het eerste longitudinale onderzoek betrof de ontwikkeling van probleemgedrag bij kinderen. Informatie was beschikbaar voor meer dan 10.000 kinderen op leeftijd 3 en 7, voor 6000 kinderen op leeftijd 10 en voor 3000 kinderen op leeftijd 12. Deze informatie werd verkregen door ouders via vragenlijsten het gedrag van hun kinderen te laten beoordelen. Ten tweede is de ontwikkeling van cognitieve vaardigheden bestudeerd. Bij 400 kinderen zijn de cognitieve vaardigheden (IQ) gemeten op vier leeftijden 5, 7, 10, en 12. Een extra maat voor cognitie is verzameld in de vorm van de CITO toets. De uitslag van de CITO toets is verzameld in een groep van 1495 kinderen. Het cross-sectionele onderzoek betrof een kleinere steekproef van de twaalfjarige kinderen waarbij speeksel verzameld was voor de bepaling van cortisol gehalten over de dag.

Het longitudinale karakter van de data biedt de mogelijkheid om de ontwikkeling van cognitief functioneren en gedragsproblemen in kaart te brengen en inzicht te verkrijgen in de oorzaken van stabiliteit en verandering in deze fenotypen tijdens de ontwikkelingen van jonge kinderen (3 jaar) tot kinderen in de prepuberale/puberale leeftijd (12 jaar).

Voor het bestuderen van deze oorzaken is gebruik gemaakt van het zogenaamde klassieke tweelingen design. In dit design wordt de overeenkomst tussen kinderen van monozygote, eeneiige, (MZ) tweelingen vergeleken met de overeenkomst tussen kinderen van dizygote, twee-eiige, (DZ) tweelingen. MZ tweelingen ontstaan als een bevruchte eicel zich in tweeën splitst. MZ tweelingen zijn genetisch identiek en dus ook altijd van hetzelfde geslacht. DZ tweelingen ontstaan na een dubbele ovulatie bij de moeder en zijn genetisch gezien niet meer verwant dan gewone broertjes of zusjes, dat wil zeggen dat ze gemiddeld 50% van hun genetisch materiaal gemeenschappelijk hebben.

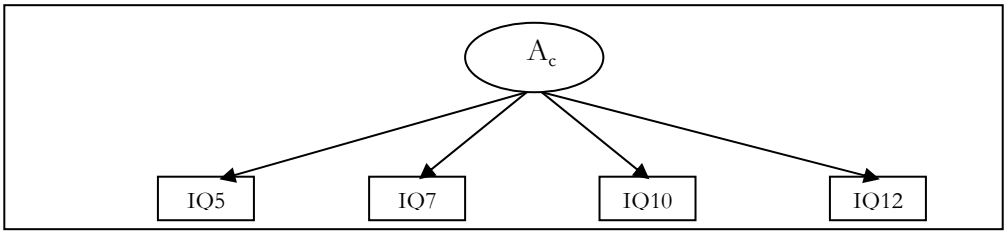
Een grotere overeenkomst tussen kinderen van MZ dan DZ tweelingen voor een bepaalde eigenschap is een eerste indicatie dat individuele verschillen in deze eigenschap mede worden bepaald door erfelijke aanleg (heritability, A). Naast deze genetisch invloed kunnen ook invloeden uit de omgeving een rol spelen. De invloeden uit de omgeving worden onderverdeeld in omgevingsinvloeden die gedeeld worden door kinderen uit een zelfde gezin (gedeelde omgeving; common environment, C) en omgevingsinvloeden die uniek zijn voor het individu (unieke omgeving; nonshared environment, E). De mate waarin MZ tweelingen en DZ tweelingen op elkaar lijken geeft informatie over het relatieve belang van A, C en E.

MZ tweelingen die in hetzelfde gezin opgroeien zijn genetisch identiek en delen dezelfde gezinsomgeving. De overeenkomst tussen MZ tweelingen is dus een functie van A + C; de invloed van erfelijke aanleg plus de invloed van de gezinsomgeving. Verschillen tussen twee kinderen van een MZ tweeling worden verklaard door invloeden die zij niet delen; de unieke omgevingsinvloeden (E). Voor DZ tweelingen die samen opgroeien geldt

ook dat zij de gezinsomgeving delen. Zij delen echter gemiddeld maar de helft van hun genetisch materiaal. De overeenkomst in DZ tweelingen is dus een functie van ½A + C.

Op grond van het patroon van MZ en DZ correlaties kan het relatieve belang van erfelijke en omgevingsinvloeden geschat worden. Als de twee kinderen van een tweeling paar (zowel MZ als DZ) niet op elkaar lijken dan speelt bij de eigenschap noch erfelijke aanleg noch hun gemeenschappelijke omgeving een rol. De individuele verschillen in gedrag of cognitie worden dan bepaald door unieke omgevingsinvloeden zoals vrienden, unieke aspecten in de relatie met de ouders, vrije tijdsbesteding etc. Als MZ en DZ tweelingen evenveel op elkaar lijken dan wordt de bestudeerde eigenschap waarschijnlijk beïnvloed door de gedeelde omgeving (zoals eetgewoonten binnen het gezin, de buurt, de school etc) en niet door hun genetische verwantschap. Wanneer de gelijkenis tussen MZ tweelingen groter is dan tussen DZ tweelingen dan is dat een sterke aanwijzingen dat genetische factoren van invloed zijn.

Figuur I. *Het 'common factor' model*



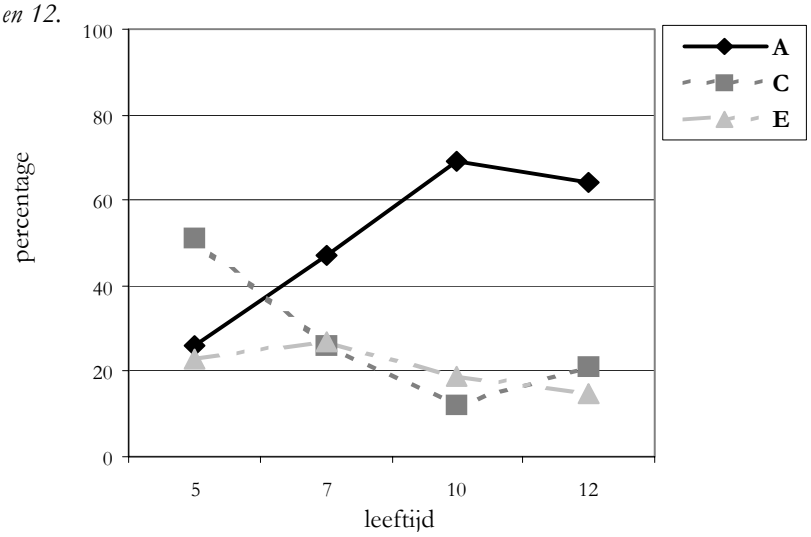
Het longitudinale onderzoek naar de oorzaken van individuele verschillen in cognitieve vaardigheden tijdens de ontwikkeling van leeftijd 5 tot 12.

Cognitieve vaardigheden zijn erg stabiel over de tijd, wat weergegeven wordt door een sterke samenhang tussen cognitie gemeten in hetzelfde kind op verschillende leeftijden. Deze samenhang kan worden weergeven door middel van een correlatie coëfficiënt (r=een getal tussen 0 en 1, waarbij 1 een perfecte samenhang weergeeft en 0 de afwezigheid van samenhang). De correlaties voor cognitieve vaardigheden over tijd zijn ongeveer .6-.7 voor tijdsintervallen van 2 tot 7 jaar. Stabiliteit in cognitieve vaardigheden blijkt voornamelijk veroorzaakt te worden door genetische factoren. Tweeënzeventig procent van de samenhang over tijd kan verklaard worden door genetische invloeden (A). De overgebleven 28% wordt veroorzaakt door omgevingsinvloeden die door beide kinderen van het tweelingpaar gedeeld worden (C). Uit de resultaten blijkt verder dat het dezelfde groep genen is die cognitieve vaardigheden op verschillende leeftijden beïnvloedt, een zogenaamde common factor structuur (zie figuur I). Echter het belang van deze groep genen neemt toe naarmate de kinderen ouder worden (zie Figuur II). De

omgevingsinvloeden die gedeeld worden door beide kinderen van een tweelingpaar, de zogenaamde gedeelde omgevingsinvloeden, worden minder belangrijk naarmate een kind ouder wordt. Het ontwikkelingspatroon voor deze gedeelde omgeving invloeden kan ook het beste beschreven worden met een 'common factor' structuur. Naast deze constante invloed van dezelfde gedeelde omgevingsfactoren blijken ook leeftijdsspecifieke gedeelde omgevingsfactoren een rol te spelen. Aangezien deze omgevingsfactoren alleen belangrijk zijn op een bepaalde leeftijd dragen ze bij aan veranderingen in cognitief functioneren over de tijd. Omgevingsinvloeden die uniek zijn voor ieder kind vertonen geen continuïteit en dragen alleen bij aan verandering van cognitieve vaardigheden over de tijd.

Figuur II.

Grafische representatie van de bronnen voor individuele verschillen in cognitieve vaardigheden op leeftijd 5, 7, 10 en 12.



In figuur II is te zien dat er sprake is van een toename van genetische invloeden op cognitie naarmate men ouder wordt. Op leeftijd 5 wordt 26% van de individuele verschillen in cognitief functioneren verklaard door verschillen in genetische aanleg. Echter, op leeftijd 10 wordt 69% van de verschillen in cognitief functioneren verklaard door genetische factoren.

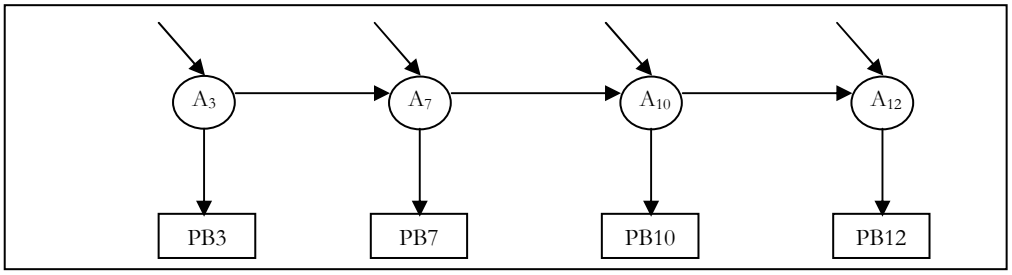
In het onderzoek is een sterke samenhang gevonden tussen cognitief functioneren op leeftijd 5, 7, 10 en 12 en de uitslag van de CITO toets. Deze samenhang van ongeveer .5 wordt verklaard door genetische factoren (A) en gedeelde omgevingsinvloeden (C). Individuele verschillen in de uitslag van de CITO toets blijken voor 60% verklaart te kunnen worden door verschillen op genetisch niveau. Significante invloeden van gedeelde

omgeving verklaren 25% van de totale variantie. De overgebleven 15% kan toegeschreven worden aan omgevingsinvloeden uniek voor ieder individu.

Het longitudinale onderzoek naar de oorzaken van individuele verschillen in gedragsproblemen in kinderen tijdens de ontwikkeling van leeftijd 3 tot 12.

De prevalentie van gedragsproblemen in Nederlandse kinderen in de leeftijd van 3 tot 12 jaar is erg hoog (5 tot 15%). De prevalentie is afhankelijk van het soort gedragsproblemen, de beoordelaar van de gedragsproblemen, en de leeftijd en geslacht van het kind. Internaliserende gedragsproblemen (angstig/depressief gedrag, teruggetrokken gedrag) komen meer voor bij meisjes terwijl Externaliserende gedragsproblemen (agressief gedrag, regelbrekend gedrag) meer voorkomen bij jongens. Wanneer ouders het gedrag van hun kinderen beoordelen rapporteren moeders meer gedragsproblemen dan vaders. De prevalentie van gedragsproblemen, gerapporteerd door de ouders, neemt af naarmate het kind ouder wordt.

Figuur III. *Het 'simplex' model*

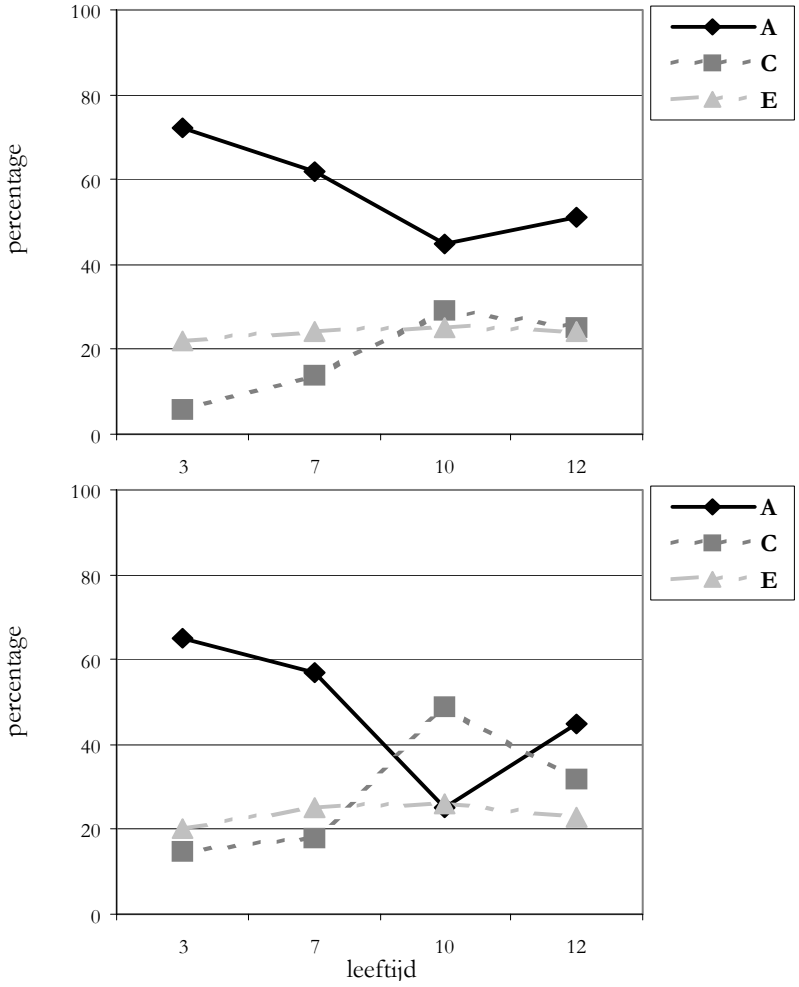


Net als voor cognitieve vaardigheden, is er gevonden dat Internaliserende en Externaliserende gedragsproblemen redelijk stabiel zijn over de tijd. De correlaties voor Internaliserende gedragsproblemen over tijd zijn: $r_{(3-7)}$=.37; $r_{(3-10)}$=.33; $r_{(3-12)}$=.30. De correlaties voor Externaliserende gedragsproblemen over tijd zijn: $r_{(3-7)}$=.55; $r_{(3-10)}$=.49; $r_{(3-12)}$=.48, waarbij de getallen tussen haakjes de leeftijdsintervallen weergeven.

Stabiliteit in gedragsproblemen kan verklaard worden door genetisch en omgevingsinvloeden. Stabiliteit in Internaliserende gedragsproblemen in jongens wordt voor 65%, 26%, en 9% verklaard door genetische (A), gedeelde omgevings (C) en unieke omgevingsinvloeden (E). Zevenenveertig procent van de stabiliteit in Internaliserende problemen in meisjes kan verklaard worden door verschillen op genetisch niveau. Gedeelde omgevingsinvloeden en unieke omgevingsinvloeden verklaren 43% en 10% van de stabiliteit in Internaliserende gedragsproblemen over tijd.

Figuur IV.

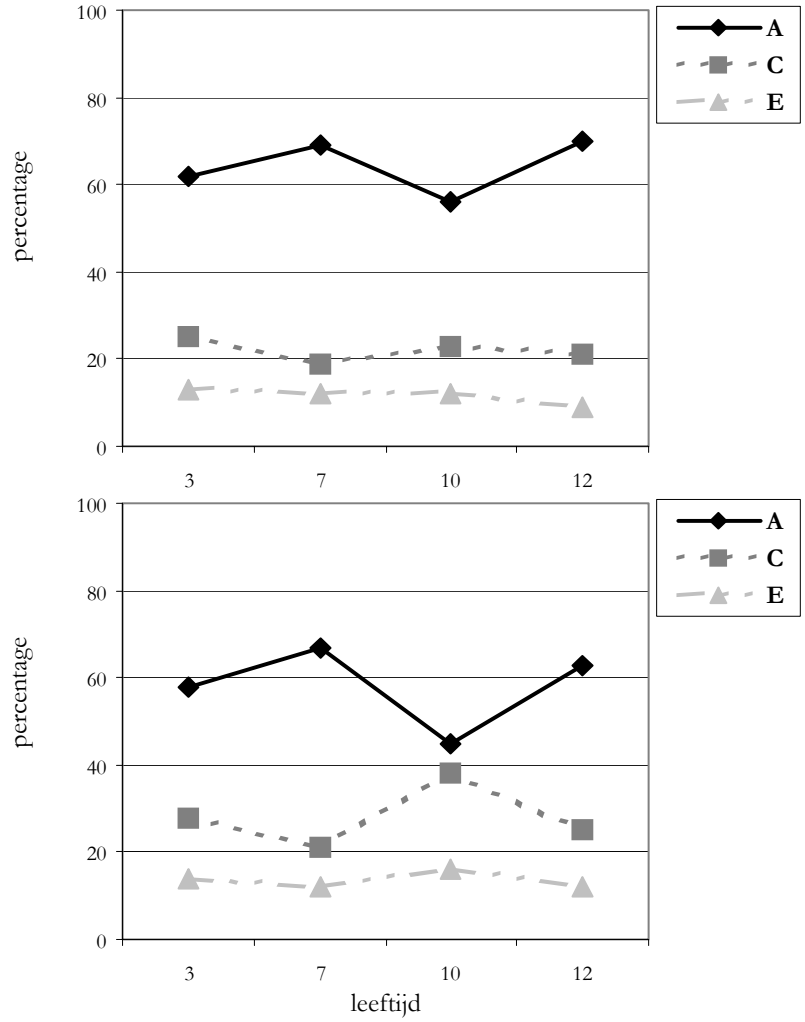
Grafische representatie van de bronnen voor individuele verschillen in Internaliserende gedragsproblemen in jongens (bovenste figuur) en meisjes (onderste figuur) op leeftijd 3, 7, 10 en 12.



Genetische invloeden blijken de belangrijkste verklarende factor voor stabiliteit in Externaliserende gedragsproblemen in jongens (73%). Negentien procent van de stabiliteit wordt verklaard door gedeelde omgevingsinvloeden en slechts 5% van de stabiliteit wordt verklaard door omgevingsfactoren uniek voor ieder individu. Voor stabiliteit in Externaliserende gedragsproblemen in meisjes blijken zowel genetische factoren als gedeelde omgevingsinvloeden een rol van betekenis te spelen. Deze factoren verklaren respectievelijk 62% en 31% van de stabiliteit over de tijd. De resterende 7% wordt verklaard door unieke omgevingsfactoren.

Figuur V.

Grafische representatie van de bronnen voor individuele verschillen in Externaliserende gedragsproblemen in jongens (bovenste figuur) en meisjes (onderste figuur) op leeftijd 3, 7, 10 en 12.



Het ontwikkelingspatroon van genetische en omgevingsinvloeden op Internaliserende en Externaliserende gedragsproblemen wordt het best beschreven met een zogenaamd simplex model (zie figuur III). Dit wil zeggen dat invloeden worden doorgegeven van en bepaalde leeftijd naar de volgende leeftijd. Naast dit transmissie proces komen er op iedere leeftijd nieuwe invloeden (innovaties) tot expressie.

Een afname van genetische invloeden en een toename van gedeelde omgevingsinvloeden op Internaliserende gedragsproblemen is zichtbaar in figuur IV. Voor genetische en omgevingsinvloeden op Externaliserende gedragsproblemen is minder verandering over de tijd zichtbaar. (Zie figuur V). Voor zowel Internaliserende als Externaliserende gedragsproblemen blijken de invloeden van unieke omgevingsinvloeden stabiel te blijven over tijd.

Conclusie van de longitudinale onderzoeken

Uit beide longitudinale onderzoeken kan geconcludeerd worden dat de ontwikkelingsprocessen van genetische en omgevingsfactoren op de verschillende fenotypes (cognitieve vaardigheden en gedragsproblemen) een verschillende patroon vertonen. Terwijl een zelfde groep genen cognitieve vaardigheden van leeftijd 5 tot en met leeftijd 12 beïnvloedt, beschrijft een transmissie proces, met genetische innovaties, de genetische invloeden op gedragsproblemen in kinderen tijdens de ontwikkeling van 3 naar 12 jaar. Verder blijkt de sterkte van de genetische en omgevingsinvloeden te veranderen over de jaren, met een toename van genetische invloeden op cognitieve vaardigheden, een afname van genetische invloeden op Internaliserende gedragsproblemen en relatief stabiele genetische invloeden op Externaliserende gedragsproblemen.

De beoordeling van gedragsproblemen

Een betrouwbare en bruikbare methode voor het verzamelen van informatie over het gedrag van kinderen is aan de hand van gestandaardiseerde vragenlijsten, waarmee ouders het gedrag van hun kind kunnen beoordelen. De overeenstemming tussen moeders en vaders over het gerapporteerde probleemgedrag is ongeveer .6. De hoogte van het verband suggereert dat ouders in staat zijn het gedrag van hun kind met enige betrouwbaarheid te beoordelen. De overeenstemming is echter niet perfect. Uit dit onderzoek blijkt dat verschillen tussen beoordelingen door moeders en vaders niet alleen het resultaat zijn van onbetrouwbaarheid van meten of enige vorm van 'rater bias' (b.v. het overschatten of onderschatten van bepaald gedrag, het hanteren van verschillende normatieve standaarden, het gebruik van een bepaalde style van antwoorden), maar dat iedere ouder, vanuit zijn eigen perspectief, informatie over het gedrag van zijn/haar kind verschaft.

Deze specifieke informatie van moeders en vaders blijkt geen continuïteit over de tijd te vertonen. Hieruit kan geconcludeerd worden dat deze ouder-specifieke informatie belangrijk is voor het bestuderen van gedragsproblemen op cross-sectioneel niveau, maar niet van significante waarde is voor het bestuderen van stabiliteit in de ontwikkeling van gedragsproblemen over de tijd.

Het blijkt dat 20% van de stabiliteit die geobserveerd wordt voor Internaliserende en Externaliserend gedrag het gevolg is van de eerder genoemde 'rater bias'. Met andere woorden, een deel van de geobserveerde stabiliteit wordt veroorzaakt door karakteristieken van de beoordelaar in plaats van stabiliteit in gedrag van het kind. De werkelijke stabiliteit in probleem gedrag is dus lager dan op grond van longitudinale studies met één beoordelaar gevonden zal worden. Deze bevinding benadrukt het belang van het gebruik van meerdere beoordelaars in het onderzoeken naar de oorzaken van individuele verschillen in de ontwikkeling van gedragproblemen.

Cortisol

Cortisol is een hormoon dat wordt afgescheiden door de bijnieren. Cortisol is vooral bekend door de rol die het speelt in de reactie op fysieke en emotionele stress. Het is echter mogelijk dat cortisol ook van invloed is op gedrag en al een rol van betekenis speelt in de vroege kinderjaren. Doordat cortisol in de prenatale of postnatale fase van invloed zou kunnen zijn op de ontwikkeling van de hersenen, zouden individuele verschillen in het cortisol gehalte mogelijk een deel van de individuele verschillen in gedragsproblemen en cognitieve vaardigheden kunnen verklaren. In een dergelijk geval wordt cortisol beschouwd als een *endophenotype* voor het onderzoeken van individuele verschillen in cognitieve vaardigheden en gedragsproblemen. Eerst moet aangetoond worden dat cortisol gehalte in verband staat met cognitie en gedrag, dat beide beïnvloed worden door genetische factoren en dat de genen die een rol lijken te spelen in cortisol gehalte tevens een rol spelen in cognitief functioneren en gedragsprobleem

Eerdere studies tonen aan dat er een verband bestaat tussen cortisol en cognitie of tussen cortisol en gedragsproblemen. Verder is het inmiddels algemeen bekend dat genetische invloeden een rol spelen in zowel cognitief functioneren als gedragsproblemen. Er is echter slechts weinig onderzoek gedaan naar de oorzaken van individuele verschillen in cortisol gehalte en de onderzoeken die zijn gedaan zijn uitgevoerd in volwassenen. Het was dus belangrijk om eerst meer inzicht te verkrijgen in de etiologie van de individuele verschillen in cortisol gehalte in kinderen.

Voor het bepalen van cortisol gehalte in kinderen is vier keer per dag op twee verschillende dagen speeksel verzameld. Resultaten tonen aan dat genen een belangrijke rol spelen in de variantie in cortisol gehalte. De erfelijkheid was gelijk voor jongens en meisjes en was het hoogst (60%) voor het cortisol gehalte in de ochtend, ongeveer 45 minuten na het opstaan.

In dit onderzoek is tot op heden echter geen significante samenhang gevonden tussen cortisol gehalte op leeftijd 12 en gedragsproblemen op leeftijd 3, 7, 10 of 12 en cognitieve vaardigheden op leeftijd 5, 7, 10 en 12. Een vervolg onderzoek zal nodig zijn om de aan of

afwezigheid van deze samenhang nader te onderzoeken en cortisol als endophenotype te gebruiken in de zoektocht naar genen die verschillen tussen kinderen in cognitief functioneren en gedragsproblemen kunnen verklaren.

Algemene conclusie

Gebaseerd op de resultaten van dit onderzoek, zoals beschreven in dit proefschrift, kan geconcludeerd worden dat genetische invloeden belangrijk zijn tijdens de ontwikkeling van kinderen. Genen beïnvloeden cognitief functioneren, gedragsproblemen en cortisol gehalte. Genen blijken ook erg belangrijk te zijn voor zowel stabiliteit als verandering van cognitie en gedragsproblemen tijdens de ontwikkeling.

Dit onderzoek laat zien dat gedraggenetische studies niet alleen belangrijk zijn voor het schatten van erfelijke en omgevingsinvloeden, maar dat dit soort studies een belangrijke aanvulling zijn op fenotypische studies bij het in kaart brengen van de ontwikkeling van jonge kinderen tot adolescenten. Het is bijvoorbeeld mogelijk om naast informatie over genetische invloeden ook inzicht te verkrijgen in de wijze waarop omgevingsinvloeden ontwikkeling beïnvloeden. Daarnaast leveren deze studies informatie op over hoe om te gaan met verschillen en overeenkomsten tussen ouderlijke beoordelingen van probleem gedrag.

References

Aardal, E., and Holm, A. C. (1995). Cortisol in saliva--reference ranges and relation to cortisol in serum. *European Journal of Clinical Chemistry and Clinical Biochemistry*, 33(12), 927-932.

Achenbach, T.M., McConaughy, S.H., and Howell, C.T. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213-232.

Achenbach, T.M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T.M. (1992). *Manual for the Child Behavior Checklist/2-3 and 1992 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.

Alarcón, M., Plomin, R., Fulker, D.W., Corley, R. and DeFries, J.C., (1998). Multivariate path analysis of specific cognitive abilities data at 12 years of age in the Colorado Adoption Project. *Behavior Genetics*, 28, 255-264.

Alarcón, M., Plomin, R., Fulker, D.W., Corley, R. and DeFries, J.C., (1999). Molarity not Modularity: Multivariate Genetic Analysis of Specific Cognitive Abilities in Parents and their 16-Year-Old Children in the Colorado Adoption Project. *Cognitive Development*, 14, 175-193.

Allanson JE, Upadhyaya M, Watson GH, Partington M, MacKenzie A, Lahey D, MacLeod H, Sarfarazi M, Broadhead W, Harper PS, et al. (1991). Watson syndrome: is it a subtype of type 1 neurofibromatosis? *Journal of Medical Genetics*, 28(11), 752-756.

Amara, S.G., Kuhar, M.J. (1993). Neurotransmitter transporters: recent progress. *Annual Review of Neuroscience*, 16, 73-93.

Anders, T.F. (1982). Biological rhythms in development. *Psychosomatic Medicine*, 44(1), 61-72.

Angold, A., Weissman, M.M., John, K., Merikangas, K.R., Prusoff, B.A., Wickramaratne, P., Gammon, G.D., and Warner, V. (1987). Parent and child reports of depressive symptoms in children at low and high risk of depression. *Journal of Child Psychology and Psychiatry*, 28, 901- 915.

Asher, S.R. (1990). Recent advances in the study of peer rejection. In S.R. Asher and J.D. (Eds.), *Peer Rejection in Childhood* (pp.3-14). Cambridge University Press.

Baker, L.A., Treloar, S.A., Reynolds, C.A., Heath, A.C., and Martin, N.G. (1996). Genetics of Educational Attainment in Australian Twins: Sex Differences and Secular Changes. *Behav Genet*. 26 (2): 89-102.

Barr, C.L., Xu, C., Kroft, J., Feng, Y., Wigg, K., Zai, G., Tannock, R., Schachar, R., Malone, M., Roberts, W., Nothen, M.M., Grunhage, F., Vandenbergh, D.J., Uhl, G., Sunohara, G., King, N., Kennedy, J.L. (2001). Haplotype study of three polymorphisms at the dopamine transporter locus confirm linkage to attention-

deficit/hyperactivity disorder. *Biological Psychiatry*, 49(4), 333-339.

Bartels, M., Boomsma, D.I., Rietveld, M.J.H., van Beijsterveldt, C.E.M., Hudziak, J.J., and E.J.C.G., van den Oord. (2002a). Disentangling Genetic, Environmental, and Rater Effects on Internalizing and Externalizing Problem Behavior in 10-year-old Twins. Submitted.

Bartels, M., Rietveld, M.J.H., Van Baal, G.C.M., Boomsma, D.I. (2002b) Genetic and Environmental Influences on the Development of Intelligence. *Behavior Genetics*, 32(4); 237- 249.

Bartels, M. Van den Berg, M., Sluyter, F., Boomsma, D.I., and De Geus, E.J.C. (2002c). Heritability of cortisol; Review and simultaneous analysis of twin studies. *Psychoneuroendocrinology*, in press.

Bartels, M., Hudziak, J.J., Boomsma, D.I., Rietveld, M.J.H., Van Beijsterveld, C.E.M., and Van den Oord, E.J.C.G. (2002d). When mom's and dad's do and don't agree: A study of parent ratings of Internalizing and Externalizing problem behavior in 12-year-old twins. *Journal of the American Academy of Child and Adolescent Psychiatry*, submitted.

Bartels, M., Van den Oord, E.J.C.G., Hudziak, J.J., Rietveld, M.J.H., van Beijsterveldt, C.E.M., and Boomsma, D.I. (2002e) Genetic and environmental mechanisms underlying stability and change in problem behaviors at ages 3, 7, 10, and 12. *Developmental Psychology*, submitted

Barth, J.M., and Parke, R.D. (1993). Parent-child relationship influences on children's transition to school. *Merill-Palmer Quarterly*, 39, 173-195.

Bender, B.G., Lerner, J.A., Kollasch, E. (1988). Mood and memory changes in asthmatic children receiving corticosteroids. *Journal of the American Academy of Child and Adolescent Psychiatry*, 27(6):720-5.

Benson, J.B., Cherny, S.S., Haith, M.M. and Fulker, D.W. (1993). Rapid Assessment of Infant Predictors of Adult IQ: Midtwin-Midparent Analyses. *Developmental Psychology*, 29, 434-447.

Beuschlein, F., Fassnacht, M., Klink, A., Allolio, B., and Reincke, M. (2001). ACTH-receptor expression, regulation and role in adrenocortial tumor formation. *European Journal of Endocrinoly*, 144(3), 199-206.

Bleichrodt, N., Drenth, P.J.D., Zaal, J.N., and Resing, W.C.M. (1984). *Revisie Amsterdamse Kinder Intelligentie Test* [Revised Amsterdam Child Intelligence Test]. Lisse, The Netherlands: Swets and Zeitlinger B.V.

Blizard, D.A. (1992). Analyzing Phenotypic Correlations in Studies with Selected Lines. *Behavior Genetics*, 22(1), 29-33.

Boivin D.B., Duffy, J.F., Kronauer, R.E., Czeisler, C.A. (1996). Dose-response relationship for resetting of human circadian clock by light. *Nature*, 379, 540-542.

Boomsma, D. I., and Molenaar, P. C., (1987). The genetic analysis of repeated measures. I. simplex models, *Behavior Genetics,* 17(2),111-123.

Boomsma, D.I. , Molenaar, P.C.M., and Orlebeke, J.F. (1990). Estimation of individual genetic and environmental factor scores. *Genetic Epidemiology,* 7, 83-91.

Boomsma, D.I., Orlebeke, J.F., and Van Baal, G.C.M. (1992). The Dutch twin register: growth data on weight and height. *Behavior Genetics*, **22**, 247-251.

Boomsma, D.I. (1993). Current status and Future Prospects in Twin Studies of the Development of Cognitive Abilities: Infancy to Old Age. In: Bouchard, T.J.j., Propping, P. (Ed.), *Twins as a Tool of Behavioral Genetics*. John Wiley and Sons Ltd. pp. 67-82.

Boomsma, D.I. and Van Baal, G.C.M. (1998). Genetic Influences on Childhood IQ in 5- and 7 Year-Old Dutch Twins. *Developmental Neuropsychology*, 14, 115-126.

Boomsma, D.I. (1998). Twin registers in Europe: an overview. *Twin Research* **1**, 34-51.

Boomsma, D.I., Vink, J.M., van Beijsterveldt, C.E.M., de Geus, E.J.C., Beem, A.L., Mulder, E.J.C.M., Riese, H., Willemsen, A.H.M., Bartels, M., van den Berg, M., Derks, E.M., de Graaff, S.C., Kupper, H.M., Polderman, J.C., Posthuma, D.I., Rietveld, M.J.H., Stubbe, J.H., Knol, L.I., Stroet, Th., van Baal, G.C.M. (2002). Netherlands Twin Register. in press *Twin Research.*

Boomsma, D.I., and Martin, N.G., (2002). Gene-Environment Interactions. In: *Biological Psychiatry*. byH.D. D'haenen, J.A. den Boer, and P. Willner (Eds). John Wiley and Sons, Ltd. Bouchard, T.J., Jr.(1981). Galton Lecture. Behaviour genetic studies of intelligence, yesterday and today: the long journey from plausibility to proof. *J.Biosoc. Sci.* **28**: 527-55.

Bouchard, T.J., Jr. and McGue, M.(1981). Familial studies of intelligence: a review. *Science*, 212, 1055-1059.

Bratko, D.(1996). Twin study of verbal and spatial abilities. *Personal and individual differences*, 21, 621-624.

Braungart-Rieker, J., Rende, R., Plomin, R., DeFries, J.C., and Fulker, D.W. (1995). Genetic mediation of longitudinal associations between family environment and childhood behavior problems. *Development and Psychopathology,* 36, 233-245.

Breslau, N., Davis, G.C., and Prabucki, K. (1988). Depressed mothers as informants in family history research. *Psychiatry Research*, 24, 345-349.

Bronfenbrenner, U. (1979). *The Ecology of Human Development*. Cambridge, MA: Harvard University Press.

Brunner, H.G., Nelen, M., Breakefield, X.O., Ropers, H.H., van Oost, B.A. (1993). Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. *Science*, 262(5133), 578-80.

Buss, A.H., and Plomin, R. (1975). *A temperament theory of personality development*. New York: Wiley.

Campbell, S.B., March, C.L., Pierce, E., Ewing, I.J. and Szumowski, E.K. (1991). Hard-to -manage preschool boys: Family context and the stability of externalizing behavior. *Journal of Abnormal Child Psychology*, 22,147-166.

Campbell, S.B. (1994). Behavior problems in preschool children: A review of recent research. *Journal of Child Psychology and Psychiatry*, 36, 113-149.

Cardon, L.R., Fulker, D.W., DeFries, J.C. and Plomin, R. (1992). Continuity and Change in General Cognitive Ability From 1 to 7 Years of Age. *Developmental Psychology*, 28, 1-10.

Carey, G. (1986). Sibling imitation and contrast effects. *Behavior Genetics*, 16, 319-341.

Carey, G (2002). *Human Genetics for the Social Sciences*. To be published by Sage Publications. Partly available at http://psych.colorado.edu/hgss/

Cases, O., Seif, I., Grimsby, J., Gaspar, P., Chen, K., Pournin, S., Muller, U., Aguet, M., Babinet, C., Shih, J.C., et al. Aggressive behavior and altered amounts of brain serotonin and norepinephrine in mice lacking MAOA. *Science*, 268(5218), 1763-1766.

Caspi, A., McClay, J., Moffitt, T.E., Mill, J., Martin, J., Craig, I.W., Taylor, A., and Poulton, R. (2002). Role of Genotype in the Cycle of Violoence in Maltreated Children. *Science*, 297: 851-853.

Cherny, S.S. and Cardon, L.R. (1994). General Cognitive Ability. In: DeFries, J.C., Plomin, R. and Fulker, D.W. (Eds.), *Nature and Nurture during Middle Childhood*. Blackwell, Oxford, England, Ch. 4.

Christenson, S.L., Rounds, T. and Franklin, M.J. (1992). Home-school collaboration: effects, issues, and opportunities. In S.L. Christenson and J.C. Conoley (Eds.), *Home-school collaboration: enhancing children's academic and social competence*. Washington, DC: NASP. pp 19-51.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.

Cohen, C.C.C., Van Goozen, S.H.M., Orlebeke, J.F., Buitelaar, J.K., and Cohen-Kettenis, P.T. (2002). A comparison of educational achievement in a national sample of Dutch Female twins and their matched singleton controls. *Twin Research*. 5(4), 273-276.

Collaer, A.S., and Hines, M. (1995). Human behavioral sex differences: a role of gonadal hormones during early development? *Psychological Bulletin*, 118, 55-107.

Comings, D.E., Gonzalez, N., Wu, S., Gade, R., Muhleman, D., Saucier, G., Johnson, P., Verde, R., Rosenthal, R.J., Lesieur, H.R., Rugle, L.J., Miller, W.B., MacMurray, J.P. (1999). Studies of the 48 bp repeat polymorphism of the DRD4 gene in impulsive, compulsive, addictive behaviors: Tourette syndrome, ADHD, pathological gambling, and substance abuse. *American Journal of Medical Genetics*, 88(4), 358-68.

Cook, E.H. Jr., Stein, M.A., Krasowski, M.D., Cox, N.J., Olkon, D.M., Kieffer, J.E., Leventhal, B.L. (1995). Association of attention-deficit disorder and the dopamine transporter gene. *American Journal of Human Genetics*, 56(4), 993-998.

Costello, E. J., Edelbrock, C. S., and Costello, A. J. (1985). Validity of the NIMH Diagnostic Interview Schedule for Children: a comparison between psychiatric and pediatric referrals. *Journal of Abnormal Child Psychology,* 13*,* 579-595.

Cowan, P.A., Cowan, C.P., Schultz, M.S., and Heming, G. (1994). Prebirth to preschool family factors in children's adaption to kindergarten. In R.D. Parke and S.G. Kellam (Eds.), *Exploring Family Relationships with other Social Contexts* (pp.75-114). Hillsdale, NJ: Erlbaum.

Crusio, W.E., 1999. Using spontaneous and induced mutations to dissect brain and behavior genetically. *Trends Neurosci.* 22(3), 100-102.

Curran, S., Mill, J., Tahir, E., Kent, L., Richards, S., Gould, A., Huckett, L., Sharp, J., Batten, C., Fernando, S., Ozbay, F., Yazgan, Y., Simonoff, E., Thompson, M., Taylor, E., Asherson, P. (2001). Association study of a dopamine transporter polymorphism and attention deficit hyperactivity disorder in UK and Turkish samples. *Molecular Psychiatry*, 6(4), 425-8.

Daly, G., Hawi, Z., Fitzgerald, M., Gill, M. (1999). Mapping susceptibility loci in attention deficit hyperactivity disorder: preferential transmission of parental alleles at DAT1, DBH and DRD5 to affected children. *Molecular Psychiatry*, 4(2), 192-196.

Dallman, M.F., Akana, S.F., Cascio, S.F., Darlington, D.N., Jacobson, L., Levin, N. (1987). Regulation of ACTH secretion: Variation on a theme of B. *Recent Progr. Horm. Res.* 43: 133-173.

Dawes, M.A., Dorn, L.D., Moss, H.B., Yao, J.K., Kirisci, L., Ammerman, R.T., Tarter, R.E. (1999). Hormonal and behavioral homeostasis in boys at risk for substance abuse. *Drug and Alcohol Dependence*, 55: 165-176.

DeFries, J.C., Plomin, R. and LaBuda, M.C.(1987). Genetic Stability of Cognitive Development From Childhood to Adulthood. *Developmental Psychology* **23**: 4-12.

De Kloet, E.R. (2000). Stress in the brain. *European Journal of Pharmacology*., 405, 187-198.

De Kloet, E.R., (1991). Brain corticosteroid receptor balance and homeostatic control. *Frontiers in Neuroendocrinology*, 12, 95-164

Demaray, M.K., Elliot, S.N. (1998). Teachers' judgement of students' academic functioning: A comparison of actual and predicted performances. *School Psychol Q*, 13, 8-24.

Desir, D., van Cauter, E., Golstein, J., Fang, V.S., Leclercq, R., Refetoff, S., Copinschi, G. (1980). Circadian and ultradian variations of ACTH and cortisol secretion. *Hormone*

Research, 13(4-5), 302-16.

DiLalla, L.F., Thompson, L.A., Plomin, R., Phillips, K., Fagan III, J.F. and Haith, M.M.(1990). Infant Predictors of Preschool and Adult IQ: A Study of Infant Twins and Their Parents. *Developmental Psychology*, 26, 759-769.

Dorin, R. I., Ferries, L. M., Roberts, B., Qualls, C. R., Veldhuis, J. D., and Lisansky, E. J., (1996). Assessment of stimulated and spontaneous adrenocorticotropin secretory dynamics identifies distinct components of cortisol feedback inhibition in healthy humans. *Journal of Clinical Endocrinology and Metabolism*, 81(11), 3883-3891.

Dorn, L.D., Hitt, S.F., Rotenstein, D. (1999). Biopsychological and cognitive differences in children with premature vs. on-time adrenarche. *Archives of Pediatric and Adolescent Medicine*, 153(2), 137-146.

Dressendörfer, R.A., Kirschbaum, C., Rohde, C., Stahl, F., Strasburger, C.J. (1992). Synthesis of a cortisol-biotin conjugate and evaluation as a tracer in an immunoassay for salivary cortisol measurement. *The Journal of Steroid Biochemistry and Molecular Biology*, 43,683-692.

Dufouil, C. and Alperovitch, A. (2000). Couple similarities for cognitive functioning and psychological health. *Journal of Clinical Epidemiology*, 53(6), 589-593.

Eaves, L.J. (1976). A model for sibling effects in man. *Heredity*, 36, 205-214.

Eaves, L.J., Last, K.A., Martin, N.G., and Jinks, J.L. (1977). A pregressive approach to non-additivity and genotype-environmental covariance in the analysis of human differences. *British Journal of Mathematical and Statistical Psychology*, 30, 1-42.

Eaves, L.J. (1982). The utility of twins. In: V.E. Anderson, W.A. Hauser, J.K. Penry, and C.F. Sing (Eds.), *Genetic basis of the epilepsies* (pp. 249-276). New York: Raven Press.

Eaves, L.J., (1984). The resolution of genotype x environment interaction in segregation analysis of nuclear families. *Genetic Epidemiology*, 1(3), 215-228.

Eaves, L.J., Long, J., and Heath, A.C.,(1986). A theory of developmental change in quantitative phenotypes applied to cognitive development. *Behavior Genetics,* 16**,** 143-162.

Eaves L.J., Fulker D.W., Heath A.C. (1989). The effects of social homogamy and cultural inheritance on the covariances of twins and their parents: a LISREL model. *Behavior Genetics*,.19(1),113-22.

Eaves, L. J., Silberg, J. L., Hewitt, J. K., Rutter, M., Meyer, J. M., Neale, M. C., and Pickles, A. (1993). Analyzing twin resemblance in multisymptom data: Genetic application of a latent class model for symptoms of conduct disorder in juvenile boys. *Behavior Genetics,* 23*,* 5-19.

Eaves, R.C. and Darch, C. (1990). The cognitive level test: It's relation with reading and mathematical achievement. *Psychol. Schools*, 27, 22-28.

Edelbrock, C. and Costello, A. J. (1988). Convergence between statistically derived behavior problem syndromes and child psychiatric diagnoses. *Journal of Abnormal Child Psychology,* 16*,* 219-231.

Edelbrock, C., Costello, A.M., Dulcan, M.K., Conover, N.C., and Kala, R. (1986). Parent-child agreement on child psychiatric symptoms assessed via structured interview. *Journal of Child Psychology and Psychiatry*, 27, 181-190.

Edelbrock, C., Rende, R., Plomin, R., and Thompson, L. (1995). A twin study of competence and problem behavior in childhood and early adolescence. *Journal of Child Psychology and Psychiatry*, 36, 775-785.

Egeland, B., Kalkoske, M., Gottesman, N., and Erickson, M.F. (1990). Preschool behavior problems: Stability and factors accounting for change. *Journal of Child Psychology and Psychiatry,* 31, 891-909.

Edwards, C.R.W., Benediktsson, R., Lindsay, R.S., Seckl, J.R. (1993). Dysfunction of placental glucocorticoid barrier: link between fetal environment and adult hypertension? *Lancet*, 341(8841), 355-7.

Eindtoets Basisonderwijs 2002, Arnhem: Citogroep, 2002.

Falconer, D.S., and Mackay, T.F.C., (1996). *Introduction to Quantitative Genetics*. Essex: Longman Scientific and Technical.

Faraone, S.V., Biederman, J., Weiffenbach, B., Keith, T., Chu, M.P., Weaver, A., Spencer, T.J., Wilens, T.E., Frazier, J., Cleves, M., Sakai, J. (1999). Dopamine D4 gene 7-repeat allele and attention deficit hyperactivity disorder. *American Journal of Psychiatry*, 156(5), 768-70.

Faraone, S.V., Doyle, A.E., Mick, E., Biederman, J. (2001). Meta-analysis of the association between the 7-repeat allele of the dopamine D(4) receptor gene and attention deficit hyperactivity disorder. *American Journal of Psychiatry*, 158(7), 1052-1057.

Ferdinand, R. F., Stijnen, T., Verhulst, F. C., and Van der Reijden M. (1999). Associations between behavioural and emotional problems in adolescence and maladjustment in young adulthood. *Journal of Adolescence,* 22*,* 123-136.

Fergusson, D.M., and Horwood, L.J. (1987). The trait and method components of ratings of conduct disorder-Part II. Factors related to the trait component of conduct disorder scores. *Journal of Child Psychology and Psychiatry*, 28, 261-272.

Ferner, R.E., Hughes, R.A., Weinman, J. (1996). Intellectual impairment in neurofibromatosis 1. *Journal of Neurological Sciences*, 138(1-2), 125-33.

Fischer, R.A., (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Trans Royal Soc Edinburgh*, 52, 915-433.

Fischman, A.J., Kastin, A.J., Graf, M.V., Moldow, R.L. (1988). Constant ligth and dark

affect the circadian rhythm of the hypothalamic-pituitary-adrenal axis. *Neuroendocrinology*, 47, 309-316.

Froehlich, J.C., Zink, R.W., Li, T.K., Christian, J.C. (2000). Analysis of heritability of hormonal responses to alcohol in twins: beta-endorphin as a potential biomarker of genetic risk for alcoholism. *Alcoholism, Clinical and Experimental Research*, 24, 265-277

Fulker, D.W. (1988). Genetic and cultural transmission in human behavior. In B.S. Weir, E.J. Eisen, M.M., Goodman, and G. Namkoong (Eds.), *Proceedings of the second international conference on quantitative genetics* (pp.318-340). Sunderland, M.A.: Sinauer.

Fulker, D.W., DeFries, J.C. and Plomin, R.(1988). Genetic influence on general mental ability increases between infancy and middle childhood. *Nature*, 336: 767-769.

Fulker, D.W., Cherny, S.S. and Cardon, L.R. (1993). Continuity and Change in Cognitive development. In: Plomin, R. and McClearn, G.E. (Eds.), *Nature, nurture and psychology*. American Psychological Association, Washington, pp. 77-97.

Fulker, D.W., and Cardon, L.R. (1993a). What Can Twins Tell Us abou the Structure and Correlates of Cognitive Abilities. In: Bouchard, T.J. jr.,Propping, P. (Eds). *Twins as a Tool of Behavioral Genetics*. John Wiley & Sons Ltd.

Gantz, I., Tashiro, T., Barcroft, C., Konda, Y., Shimoto, Y., Miwa, H., Glover, T., Munzert, G., Yamada, T. (1993). Localization of the genes encoding the melanocortin-2(adrenocorticotropic hormone) and melanocortin-3 receptors to chromosomes 18p11.2 and 20q13.2-q13.3 by fluorescence in situ hybridization. *Genomics*, 18, 166-167.

García, F.J. and Rosel, J. (2001). Family and personal correlates of academic achievement. *Psychological Reports*, 88, 533-547.

Geldermans, A. (2001). Toetsgekte, of duidelijkheid? J/M September, 52-56.

Gerra, G., Zaimovic, A., Giucastro, G., Folli, F., Tessoni, A., Avanzini, P., Caccavari, R., Bernasconi, S., Brambilla, F. (1998). Neurotransmitter-Hormonal Responses to Psychological Stress in Peripubertal Subjects; Relationship to Aggressive Behavior. *Life Sciences*, 62 (7), 617 -625.

Ghodsian, M., Fogelman, K., Lambert, L., and Tibbenham, A. (1980). Changes in behavior ratings of a national sample of children. *British Journal of Social and Clinical Psychology*, 19, 247-256.

Gill, M., Daly, G., Heron, S., Hawi, Z., Fitzgerald, M. (1997). Confirmation of association between attention deficit hyperactivity disorder and a dopamine transporter polymorphism. *Molecular Psychiatry*, (4), 311-313.

Gjone, H., and Stevenson, J. (1997). The association between Internalizing and Externalizing behavior in childhood and early adolescence: Genetic and environmental common influences? *Journal of Abnormal Child Psychology*, 25, 277-286.

Glascoe, F.P. (2001). Can Teachers' Global Ratings Indentify Children with Academic Problems? *Developmental and Behavioral Pediatrics*, 22(3), 163-168.

Goodman, R., and Stevenson, J. (1991). Parental criticism and warmth towards unrecognized monozygotic twins. *Behavior* and *Brain Sciences*, 14: 394-395.

Goodyer, I. M., Herbert, J., Tamplin, A., and Altham, P. M. (2000). Recent life events, cortisol, dehydroepiandrosterone and the onset of major depression in high-risk adolescents. *British Journal of Psychiatry*, 177, 499-504.

Gould, E. (1994). The effects of adrenal steroids and excitatory input on neuronal birth and survival. *Ann N Y Acad Sci*, 743, 73-92.

Graham, P., and Rutter, M. (1973). Psychiatric disorder in the young adolescent. *Proceedings of the Royal Society of Medicine*, 66, 1226-1229.

Granger, D.A., Weisz, J.R., Kauneckis, D. (1994). Neuroendocrine reactivity, internalizing behavior problems, and control-related cognition in clinic-referred children and adolescents. *Journal of Abnormal Psychology*, 103(2), 267-276.

Granger, D.A., Weisz, J.R., McCracken, J.T., Ikeda, S.C., Douglas, P. (1996). Reciprocal influences among adrenocortical activation, psychosocial processes, and the behavioral adjustment of clinic-referred children. *Child Development*, 67(6), 3250-3262.

Guttman, L. (1954). A New Approach to Factor Analysis: The Radax. In: Lazarsfeld, P.F. (Ed.), *Mathematical thinking in the social sciences*. Free Press, Glencoe Ill, pp. 258-349.

Harris, J.R. (1995). Where's the child's environment? A group socialization theory of development. *Psychological Review*, 102, 458-489.

Heath, A.C., and Eaves, L.J. (1985). Resolving the effects of phenotype and social background on mate selection. *Behavior Genetics*, 15(1), 15-30.

Heath, A. C., Madden, P. F., Reich, W., Bucholz, K. K., Slutske, W., Bierut, L. J., Neuman, R. J., and Todd, R. D. (1998). Latent class and factor analysis of DSM-IV ADHD: a twin study of female adolescents. *Journal of the Amercian Academy Child Adolesceny Psychiatry,* 37*,* 848-857.

Heath, A.C., Todorov, A.A., Nelson, E.C., Madden, P.A., Bucholz, K.K., and Martin, N.G. (2002). Gene-environment interaction effects on behavioral variation and risk of complex disorders: the example of alcoholism and other psychiatric disorders. *Twin Research*, 5(1), 30-37.

Heim, C., Newport, D. J., Heit, S., Graham, Y. P., Wilcox, M., Bonsall, R., Miller, A. H., and Nemeroff, C. B. (2000). Pituitary-adrenal and autonomic responses to stress in women after sexual and physical abuse in childhood. *The Journal of the American Medical Association*, 284(5), 592-597.

Helgeson, V.S. (2002). *The psychology of gender*. Pearson Education, Inc. Upper Saddle River, New Jersey.

Henderson AS, Easteal S, Jorm AF, Mackinnon AJ, Korten AE, Christensen H, Croft L, Jacomb PA. (1995). Apolipoprotein E allele epsilon 4, dementia, and cognitive decline in a population sample. *Lancet*, 346(8987), 1387-1390.

Herjanie, B., and Reich, W. (1982). Development of a structured psychiatric interview for children:Agreement between child and parent on individual symptoms. *Journal of Abnormal Child Psychology*, 10, 307-324.

Heuser, I., von Bardeleben, U., Boll, E., Holsboer, F. (1988). Response of ACTH and cortisol to human corticotropin-releasing hormone after short-term abstention from alcohol abuse. *Biological Psychiatry*, 24(3), 316-21.

Hewitt, J.K., Silberg, J.L., Neale, M.C., Eaves, L.J., and Erickson, M., (1992). The analysis of parental ratings of children's behavior using LISREL. *Behavior Genetics*, 22, 293-317.

Hofman, K.J., Harris, E.L., Bryan, R.N., Denckla, M.B. (1994). Neurofibromatosis type 1: the cognitive phenotype. *Journal of Pediatrics*, 24(4), S1-8.

Holmes, J., Payton, A., Barrett, J.H., Hever, T., Fitzpatrick, H., Trumper, A.L., Harrington, R., McGuffin, P., Owen, M., Ollier, W., Worthington, J., Thapar, A. (2000). A family-based and case-control association study of the dopamine D4 receptor gene and dopamine transporter gene in attention deficit hyperactivity disorder. *Molecular Psychiatry*, 5(5), 523-530.

Holsboer, F., and Barden, N. (1996). Antidepressants and hypothalamic pituitary adrenocortical regulation. *Endocrine Reviews*, 17, 187-205.

Hudziak, J.J., Heath, A.C., Madden, P.F., Reich, W., Bucholz, K.K., Slutkske, W., Bierut, L.J., Neuman, R.J., and Todd, R.D. (1998). Latent class and factor analysis of DSM-IV ADHD: a twin study of female adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 37, 848-857.

Hudziak, J.J., Rudiger, L.P., Neale, M.C., Heath, A.C., and Todd, R.D. (2000). A twin study of Inattentive, Aggressive and Anxious/Depressed Behaviors. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39(4), 469-476.

Huizenga, N.A., Koper, J.W., De Lange, P., Pols, H.A., Stolk, R.P., Burger, H., Grobbee, D.E., Brinkmann, A.O., De Jong, F.H., Lamberts, S.W. (1998). A polymorphism in the glucocorticoid receptor gene may be associated with and increased sensitivity to glucocorticoids in vivo. *Journal of Clinical Endocrinology and Metabolism*, 83, 144-151.

Humphreys, L.G., and Davey, T.C.(1988). Continuity in Intellectual Growth from 12 Months to 9 Years. *Intelligence*, 12, 183-197.

Inglis, G. C., Ingram, M. C., Holloway, C. D., Swan, L., Birnie, D., Hillis, W. S., Davies, E., Fraser, R., and Connell, J. M. (1999). Familial pattern of corticosteroids and their metabolism in adult human subjects--the Scottish Adult Twin Study. *Journal of Clinical Endocrinology and Metabolism*, 84(11), 4132-4137.

Jacobson, L., Sapolsky, R., (1991). The role of the hippocampus in feedback regulation of the hypothalamic-pituitary-adrenocortical axis. *Endocrine Reviews*, 12, 118-134.

Jensen A.R. (1972). *Genetics and Education*, Harper and Row, New York.

Jensen, A.R. (1978). Genetic and behavioral effects of nonrandom mating. In Osborne, R.T., Nolbe, C.E., and Weyl, N. (Eds.), *Human Variation: The Biopsychology of Age, Race and Sex,* Academic Press, New York.

Jensen, P.S., Traylor, J., Xenakis, S.N., and Davis, H. (1988). Child psychopathology rating scales and interrater agreement I: Parents' gender and psychiatric symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry*, 27(6), 798-801..

Jöreskog, K.G., and Sörbom D.(1996). PRELIS2: User's Reference Guide. Scientific Software International, Inc. Chicago.

Kagan, J., Reznick, J.S., Snidman, N. (1987). The physiology and psychology of behavioral inhibition in children. *Child Development*, 58, 1459-1473.

Karl, M., and Schulte, H.M. (1994). Familiäre Glucocorticoidresistenz als Differentialdiagnose des Hypercortisolismus. Klinische und molekulargenetische Aspekte. [Familial glucocorticoid resistance as the differential diagnosis of hypercortisolism. Clinical and molecular genetic aspects]. *Dtsch Med Wochenschr*, 119, 74-79.

Kashini, J.H., Orvaschel, H., Burk, J.P. and Reid, J.C. (1985). Informant variance: The issue of parent-child disagreement. *Journal of the American Academy of Child Psychiatry*, 24, 437-441.

Kasius, M. C., Ferdinand, R. F., van den, B. H., and Verhulst, F. C. (1997). Associations between different diagnostic approaches for child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry,* 38, 625-632.

Kazdin, A.E., Esveldt-Dawson, K., Unis, A.S., and Raneurello, M.D. (1983). Child and parent evaluation of depression and aggression in psychiatric inpatient children. *Journal of Abnormal Child Psychology* **11**, 401-413.

Keenan PA, Jacobson MW, Soleymani RM, Mayes MD, Stress ME, Yaldoo DT. (1996). The effect on memory of chronic prednisone treatment in patients with systemic disease. *Neurology*, 47(6), 1396-1402.

Kendler, K.S., and Eaves, L.J. (1986). Models for the joint effect of genotype and environment on liability to psychiatric illness. *American Journal of Psychiatry*, 143, 279-289.

Kirschbaum, C., and Hellhammer, D.H. (1989). Salivary Cortisol in psychobiological research: An overview. *Neuropsychobiology*, 22, 150-169.

Kirschbaum, C., Strasburger, C.J., Jammers, W., Hellhammer, D.H., 1989. Cortisol and

behavior: 1. adaptation of a radioimmunoassay kit for reliable and inexpensive salivary cortisol determination. *Pharmacol. Biochem. Behav.* 34, 747-751

Kirschbaum, C., Wüst, S., Faig, H. G., and Hellhammer, D. H., (1992). Heritability of cortisol responses to human corticotrophin-releasing hormone, ergometry, and psychological stress in humans. *Journal of Clinical Endocrinology and Metabolism*, 75(6),1526-1530.

Kirschbaum, C., and Hellhammer, D. H., (1994). Salivary cortisol in psychoneuroendocrine research: recent developments and applications. P*sychoneuroendocrinology*, 19(4),313-333.

Kish, L. (1965). *Survey Sampling*. New York Wiley.

Koot, H.M. (1993). *Problem behaviour in Dutch pre-schoolers*. Doctoral dissertation, Erasmus University, Rotterdam.

Koot, H.M., (1995). Longitudinal studies of general population and community samples. In: *The epidemiology of child and adolescent psychopathology,* Verhulst FC, Koot HM, eds. London: Oxford University Press.

Koot, H.M., Van den Oord, E.J.C.G., Verhulst, F.C., and Boomsma, D.I., (1997). Behavioural and emotional problems in young pre-schoolers: cross-cultural testing of the validity of the Child Behavior Checklist 2/3. *Journal of Abnormal Child Psychology*, 25,183-196.

Kovacs, G.L., Fekete, M., Szabo, G., Telegdy, G. (1987). Action of the ACTH-corticosteroid axis on the central nervous system. *Frontiers of Hormone Research*, 15, 79-127.

Krueger, R.F., Moffitt, T.E., Caspi, A., Bleske, A. and Silva, P.A. (1998). Assortative mating for antisocial behavior: development and methodological implications. *Behavior Genetics*, 28(3), 173-186.

Kruesli, M.J.P., Schmidt, M.E., Donnelly, M., Hibbs, E.D., Hamburger, S.D. (1989). Urinary Free Cortisol in Output and Disruptive Behavior in Children. *Journal of the American Academy of* Child Adolescent Psychiatry, 28, 441-443.

Kuhn, C. (1989). Adrenocortical and gonadal steroids in behavioral cardiovascular medicine. In: N. Schneiderman, S.M., Weis, P.G. Kaufman (Eds). *Handbook of Research Methods on Cardiovascular Behavioral Medicine*, pp. 185-204. New York: Plenum.

LaBuda, M.C., DeFries, J.C., Plomin, R. and Fulker, D.W.(1986). Longitudinal stability of cognitive ability from infancy to early childhood: genetic and environmental etiologies. *Child Development*, 57, 1142-1150.

Ladd, G.W., and Price, J.M. (1987). Predicting children's social and school adjustment following the transition from preschool to kindergarten. *Child Development*, 58, 1168-1189.

LaHoste, G.J., Swanson, J.M., Wigal, S.B., Glabe, C., Wigal, T., King, N., Kennedy, J.L. (1996). Dopamine D4 receptor gene polymorphism is associated with attention deficit hyperactivity disorder. *Molecular Psychiatry*, 1(2), 121-124.

Lamd, M.E., and Nash, A. (1989). Infant-mother attachment, sociability and peer competence. In T.J. Berndt and G.W. Ladd (Eds). *Peer relations in development* (pp.219-246). New York: Wiley.

Lange, K., Westlake, J., and Spence, M.A. (1976). Extension to pedigree analysis: III. Variance components by the scoring method. *Annals of Human Genetics*, 48, 47-59.

Leve, L.D., Winebarger, A.A., Fagot, B.I., Reid, J.B., and Goldsmith, H.H. (1998). Environmental and genetic variance in children's observed and reported maladaptive behavior. *Child development*, 69, 1286-1298.

Levine, S. (1994). The ontogeny of the HPA axis: the influence of maternal factors. *Ann. N. Y. Acad. Sci.*, 746, 275-288.

Levitt, N.S., Lindsay, R.S., Holmes, M.C., Seckl, J.R. (1996). Dexamethasone in the last week of pregnancy attenuates hippocampal glucocorticoid receptor gene expression and elevates blood pressure in the adult offspring in the rat. *Neuroendocrinology*, 64(6), 412-8.

Li, A., Li, K.X., Marui, S., Krozowski, Z.S., Batista, M.C., Whorwood, C.B., Arnhold, I.J., Shackleton, C.H., Mendonca, B.B., Stewart, P.M. (1997). Apparent mineralocorticoid excess in a Brazilian kindred: hypertension in the heterozygote state. *Journal of Hypertension*, 15, 1397-1402.

Lichtenstein, R. (1984). Predicting school performance of preschool children from parent reports. *Journal of Abnormal Child Psychology*, 12, 79-93.

Liddle, G. (1960). Tests of pituitary-adrenal suppressibility in the diagnosis of Cushing's syndrome. *Journal of Clinical Endocrinology and Metabolism*, 20, 1539-1560.

Lindsay, R.S., Lindsay, R.M., Waddell, B.J., Seckl, J.R.. (1996a). Prenatal glucocorticoid exposure leads to offspring hyperglycaemia in the rat: studies with the 11 beta-hydroxysteroid dehydrogenase inhibitor carbenoxolone. *Diabetologia*, 39(11), 1299-1305.

Lindsay, R.S., Lindsay, R.M., Edwards, C.R., Seckl, J.R. (1996b). Inhibition of 11-beta-hydroxysteroid dehydrogenase in pregnant rats and the programming of blood pressure in the offspring. *Hypertension*, 27(6), 1200-1204.

Linkowski, P., Mendlewicz, J., Leclercq, R., Brasseur, M., Hubain, P., Golstein, J., Copinschi, G., and Van Cauter, E. (1985). The 24-hour profile of adrenocorticotropin and cortisol in major depressive illness. *Journal of Clinical Endocrinology and Metabolism*, 61(3), 429-438.

Linkowski, P., Van Onderbergen, A., Kerkhofs, M., Bosson, D., Mendlewicz, J., and Van

Cauter, E. (1993). Twin study of the 24-h cortisol profile: evidence for genetic control of the human circadian clock. *The American Journal of Physiology*, 264(2 Pt 1), E173-E181.

Loeber, R., Green, S.M., Lahey, B.B., and Stouthamer-Loeber, M. (1989). Optimal informants on childhood disruptive behaviors. *Development and Psychopathology*, 1, 317-337.

Loeber, R., Green, S.M., Lahey, B.B. and Stouthamer-Loeber, M. (1991). Differences and similarities between children, mothers, and teachers as informants on disruptive child behavior. *Journal of Abnormal Child Psychology*, 19, 75-95.

Loehlin, J.C., Horn, J.M. and Willerman, L.(1989). Modeling IQ change: evidence from the Texas Adoption Project. *Child Development*, 60, 993-1004.

Loehlin, J.C. and Nichols, R.C. (1976). *Heridity, environment and personality*. Austin: University of Texas Press.

Lopez Bernal, A., Anderson, A. B., Parry, D. M., and Turnbull, A. C. (1980). Evidence that fetal membranes are not involved in cortisol metabolism: study of dichorionic twin pregnancies. *American Journal of Obstetrics and Gynecology*, 138(8), 1168-1172.

Lopez, J.F., Chalmers, D.T., Little, K.Y., and Watson, S.J. (1998). Regulation of serotonin(1A), glucocorticoid, and mineralocorticoid receptor in rat and human hippocampus: Implications for the neurobiology of depression. *Biological Psychiatry*, 43, 547-573.

Lucki, I. (1998). The spectrum of behaviors influenced by serotonin. *Biological Psychiatry*, 44(3), 151-162.

Lupien, S.J., de Leon, M., de Santi, S., Convit, A., Tarshish, C., Nair, N.P., Thakur, M., McEwen, B.S., Hauger, R.L., Meaney, M.J. (1998). Cortisol levels during human aging predict hippocampal atrophy and memory deficits. *Nature Neuroscience*, 1(1), 69-73.

Majoribanks, K. (1994). Families, schools and children's learning: a study of children's learning environment. *International Journal of Educational Research*, 21, 439-555.

Makalowski, W., Boguski, M.S. (1998). Evolutionary parameters of the transcribed mammalian genome: an analysis of 2,820 orthologous rodent and human sequences. *Proceedings of the National Academy of Science U S A*, 95(16), 9407-12.

Mantero, F., and Boscaro, M. (1992). Glucocorticoid-dependent hypertension. *The Journal of Steroid Biochemistry and Molecular Biology*, 43(5), 409-413.

Manuck, S.B., Flory, J.D., Ferrell, R.E., Mann, J.J., Muldoon, M.F. (2000). A regulatory polymorphism of the monoamine oxidase-A gene may be associated with variability in aggression, impulsivity, and central nervous system serotonergic responsivity. *Psychiatry Research*, 95(1), 9-23.

Martin, N.G., and Martin, P.G. (1975). The inheritance of scholastic abilities in a sample of twins. *Annals of Human Genetics*, 39, 213-218.

Martin, N.G. (1975). The inheritance of scholastic achievement in a sample of twins II. Genetic analysis of examination results. *Annals of Human Genetics*, 39, 219-229.

Martin, N.G., and Eaves, L.J. (1977). The genetical analysis of covariance structure. *Heredity*, 38, 79-95.

Martin, N.G., Eaves, L.J., and Fulker, D.W. (1979) The genetical relationship of impulsiveness and sensation seeking to Eysenck's personality dimensions. *Acta Genetic Med Gemellol*, 28(3), 197-210.

Martin, N.G., Boomsma, D.I., Machin, G.A., (1997). A twin-pronged attack on complex traits. *Nature Genetics*, 17, 387-391.

Mascie-Taylor, C.G.N., and Vandenberg, S.G. (1988). Assortative mating for IQ and personality due to propinquity and personal preference. *Behavior Genetics*, 18 (3), 339-345.

Mason, J.W. (1968). A review of psychoendocrine research on the pituitary-adrenal cortical system. *Psychosomatic Medicine*, 30, 576-607.

Mather, K., and Jinks, J.L., (1977). *Introduction to Biometrical Genetics*. Chapman and Hall Ltd, London.

Mathews, C.A., and Reus, V.I. (2001). Assortative mating in the affective disorders: a systematic review and meta-analysis. *Comprehensive Psychology*, 42(4), 257-262.

Matthews, S.G. (2000). Antenatal Glucocorticoids and Programming of the Developing CNS. *Pediatric Research*, 47(3), 291-300.

Maxwell, J.D., Boyle, L.A., Greig, W.R., Buchanan, W.W., (1969). Plasma Corticosteroids in Healthy Twin Pairs. *Journal of Medical Genetics*, 6, 294-297.

McArdle, J.J., and Goldsmith, H.H. (1990). Alternative common factor models for multivariate biometric analyses. *Behavior Genetics*, 20, 569-608.

McBurnett, K., Lahey, B. B., Frick, P. J., Risch, C., Loeber, R., Hart, E. L., Christ, M. A., Hanson, K. S. (1991). Anxiety, inhibition, and conduct disorder in children: II. Relation to salivary cortisol. *Journal of the American Academy fo Child and Adolescent Psychiatry*, 30(2),192-196.

McBurnett, K., Lahey, B.B., Capasso, L., Loeber, R. (1996). Aggressive symptoms and salivary cortisol in clinic-referred boys with conduct disorder. *An N Y Acad Sc*, 794, 169-178.

McBurnett, K., Lahey, B.B., Rathouz, P.J., Loeber, R. (2000). Low salivary cortisol and persistent aggression in boys referred for disruptive behavior. *Archives of General Psychiatry*, 57(1),38-43.

McCartney, K., Harris, M.J. and Bernieri, F.(1990). Growing up and growing apart: a developmental meta-analysis of twin studies. *Psychological Bulletin*, 107, 226-237.

McEwen, B.S., 2000. The neurobiology of stress: from serendipity to clinical relevance.

Brain Res. 886, 172-189.

McGue, M., Bouchard, T.J., Iacono, W.G. and Lykken, D.T.(1993). Behavioral Genetics of Cognitive Ability: A Life-Span Perspective. In: Plomin, R. and McClearn, G.E. (Eds.), *Nature, nurture and psychology*. American Psychology Association, Washington, pp. 59-76.

Meijer, O.C., de Kloet, E.R., McEwen, B.S. (2000). Corticosteroid Receptors. in *Encyclopedia of Stress*. Academic press.

Meikle, A. W., Stringham, J. D., Woodward, M. G., and Bishop, D. T. (1988). Heritability of variation of plasma cortisol levels. *Metabolism*, 37(6), 514-517.

Mendel, G. (1865). *Experiments in Plant Hybridozation*. Available at the Mendelwed. http://www.netspace.org/Mendelweb.

Mendlewicz, J., Hoffmann, G., Kerkhofs, M., and Linkowski, P. (1984). Electroencephalogram and neuroendocrine parameters in pubertal and adolescent depressed children. A case report study. *J Affect Disord*, 6(3-4), 265-272.

Milford, D.V., Shackleton, C.H.L., Stewart, P.M. (1994). Mineralocorticoid hypertension and congenital deficiency of 11 β-hydroxysteroid dehydrogenase in a family with the syndrome of 'apparent' mineralocorticoid excess. *Clinical Endocrinology*, 43, 241-246.

Monk, C.S., and Nelson, C.A. (2002). The effects of hydrocortisone on cognitive and neural function: a behavioral and event-related potential investigation. *Neuropsychopharmacology*, 26(4):505-519.

Morley, K.I., Montgomery, G.W. (2001). The genetics of cognitive processes: candidate genes in humans and animals. *Behavior Genetics*, 31(6), 511-531.

Moss, H.B., Vanyukov, M.M., Martin, C.S. (1995). Salivary cortisol responses and the risk for substance abuse in prepubertal boys. *Biological Psychiatry*, 38(8), 547-555.

Mountjoy, K.G., Robbins, L.S., Mortrud, M.T., Cone, R.D. (1992). *Science*, 257, 48-51.

Muglia, P., Jain, U., Macciardi, F., Kennedy, J.L. (2000). Adult attention deficit hyperactivity disorder and the dopamine D4 receptor gene. *American Journal of Medical Genetics*, 96(3), 273-7.

Mullen, P.E., Martin, J.L., Anderson, J.C., Romans, S.E., Herbison, G.P. (1996). The long-term impact of the physical, emotional, and sexual abuse of children: a community study. *Child. Abuse. Negl.*, 20, 7-21.

Muthén, B., and Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of nonnormal Likert variables. *British Journal of Mathematical and Statistical Psychology*, 38, 171-189.

Nagoshi, C.T., Johnson, R.C., and Ahern, F.M. (1987). Phenotypic assortative mating vs. social homogamy among Japanese and Chinese parents in the Hawaii Family Study of Cognition. *Behavior Genetics*, 17 (5), 477-485.

Neale , M.C., and Stevenson, J. (1989). Rater bias in the EASI Temperament scales: a twin study. Journal of Personality and Social Psychology, 56, 446-455.

Neale, M.C., and Cardon, L.R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer Academic.

Neale, M.C., Boker, S.M., Xie, G., and Maes, H.H. (1999). *Mx: Statistical Modeling*. VCU Box 900126, Richmond, VA 23298: Department of Psychiatry. 5th Edition.

Neuman, R. J., Todd, R. D., Heath, A. C., Reich, W., Hudziak, J. J., Bucholz, K. K., Madden, P. A., Begleiter, H., Porjesz, B., Kuperman, S., Hesselbrock, V., and Reich, T. (1999). Evaluation of ADHD typology in three contrasting samples: a latent class approach. *Journal of the Americam Academy of Child Adolescent Psychiatry*, 38, 25-33.

Newcomer, J.W., Craft, S., Askins, K., Hershey, T., Bardgett, M.E., Csernansky, J.G., Gagliardi, A.E., Vogler, G. (1998). Glucocorticoid interactions with memory function in schizophrenia. *Psychoneuroendocrinology*, 23(1), 65-72.

Newcomer, J.W., Selke, G., Melson, A.K., Hershey, T., Craft, S., Richards, K., Alderson, A.L. (1999). Decreased memory performance in healthy humans induced by stress-level cortisol treatment. *Archives of General Psychiatry*, 56(6), 527-33.

Nichols, R.C. (1965). The National Merit Twin Study. In S.G. Vandenberg (Ed.), *Methods and goals in human behavior genetic*. pp 231-244. New York: Academic Press.

Norman, R. J., and Joubert, S. M. (1982). Fetal adrenal gland maturation in growth-retarded twin pregnancies. *Acta Genet. Med. Gemellol.* (Roma), 31(3-4), 59-62.

Norman, R. J., Deppe, W. M., Coutts, P. C., Marivate, M., and Joubert, S. M. (1983a). Twin pregnancy as a model for studies in fetal cortisol concentrations in labour: relation to prostaglandins, prolactin and ACTH. *British Journal of Obstetrics and Gynaecology*, 90(11), 1033 -1039.

Norman, R. J., Joubert, S. M., and Marivate, M. (1983b). Amniotic fluid phospholipids and glucocorticoids in multiple pregnancy. *British Journal of Obstetrics and Gynaecology*, 90(1), 51-55. Norman, R. J., Maharaj, C., Adhikari, M., and Joubert, S. M. (1984). Adrenal sensitivity to adrenocorticotrophin in twin neonates--effect of birth order and growth retardation. *S. Afr. Med. J.*, 66(3), 90-92.

North, K., Joy, P., Yuille, D., Cocks, N., Hutchins, P. (1995). Cognitive function and academic performance in children with neurofibromatosis type 1. *Developmental Medicine and Child Neurology*, 37(5), 427-36.

Nurnberger, J.I., Gershon, E.S., Simmons, S., Ebert, M., Kessler, L.R., Dibble, E.D., Jimerson, S.S., Brown, G.M., Gold, P., Jimerson, D.C., Guroff, J.J., Storch, F.I., 1982. Behavioral, biochemical and neuroendocrine responses to amphetamine in normal twins and 'well-state' bipolar patients. *Psychoneuroendocrinology* 7, 163-176.

Nurnberger, J. I., Jimerson, D. C., Simmons-Alling, S., Tamminga, C., Nadi, N. S.,

Lawrence, D., Sitaram, N., Gillin, J. C., and Gershon, E. S. (1983). Behavioral, physiological, and neuroendocrine responses to arecoline in normal twins and "well state" bipolar patients. *Psychiatry Research*, 9(3),191-200.

Nyirenda, M.J., Seckl, J.R. (1998). Intrauterine events and the programming of adulthood disease: The role of fetal glucocorticoid exposure (Review). *Int J Mol Med*, 2(5), 607-614.

O'Connor TG, Neiderhiser JM, Reiss D, Hetherington EM, and Plomin R (1998). Genetic contributions to continuity, change, and co-occurrence of antisocial and depressive symptoms in adolescence. *Journal of Child Psychology and Psychiatry*, 39, 323-336.

Okumura, T., Nakajima, Y., Takamatsu, T., Matsuoka, M. (1995). Column-switching high -performance liquid chromatographic system with a laser-induced fluorimetric detector for direct, automated assay of salivary cortisol. *J. Chromatog. B: Biomed. Applications*, 670, 11-20.

Ozonoff ,S. (1999). Cognitive impairment in neurofibromatosis type 1. *American Journal of Medical Genetics*, 89(1):45-52.

Palmer, C.G.S., Bailey, J.N., Ramsey, C., Cantwell, D., Sinsheimer, J.S., Del'Homme, M., McGough, J., Woodward, J.A., Asarnow, R., Asarnow, J., Nelson, S., Smalley, SL. (1999). No evidence of linkage or linkage disequilibrium between DAT1 and attention deficit hyperactivity disorder in a large sample. *Psychiatric Genetics*, 9(3), 157-160.

Parke, R.D., and Kellam, S.G. (Eds., 1994). *Exploring family relationships with other social contexts.* Hillsdale, NJ: Erlbaum.

Patrick, C.L. (2000). Genetic and Environmental Influences on the Development of Cognitive Abilities: Evidence From the Field of Developmental Behavior Genetics. *Journal of School Psychology*, 38, 79-108.

Pedersen, B. K., and Hoffman-Goetz, L. (2000). Exercise and the immune system: regulation, integration, and adaptation. *Physiol. Rev.*, 80(3), 1055-1081.

Peterson JL, and Zill, N. (1986) Marital disruption, parent-child relationships, and behavior problems in children. *Journal of Marriage and the Family*, 48,295-307.

Petrill, S.A., and Thompson, L.A. (1993). The Phenotypic and Genetic Relationships Among Measures of Cognitive Ability, Temperament, and Scholastic Achievement. *Behavior Genetics*, 23, 6: 511-518.

Petronis, A. (2001). Human morbid genetics revisited: relevance of epigenetics. *Trends in Genetics*, 17(3), 142-146.

Phares, V., Compas, B.E., and Howell, D.C. (1989). Perspectives on child behavior problems: Comparisons of children's self-reports with parent and teacher reports. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1, 68-71.

Philips, P.C. (1998). The language of gene interaction. *Genetics*, 149, 1167-1171.

Philips, K., Fulker, D.W., Carey, G., and Nagoski, C.T. (1988). Direct marital assortment for cognitive and personality variables. *Behavior Genetics*, 18 (3), 347-356.

Phillips, K., Fulker, D.W. (1989). Quantitative genetic analysis of longitudinal trends in adoption designs with application to IQ in the Colorado Adoption Project. *Behavior Genetics*, 19, 621-658.

Pinheiro, M.M., Liberman, B., Salgado, L.R., Goldman, J., Nery, M., Cukiert, A. (1999). Identical twins discordant for Cushing's disease: case report. *Arquivos Neuro-psiquiatria*, 57, 686-688.

Piaget, J. (1954). The construction of reality in the Child. New York. Basic Books.

Pijl, Y.J., Hofman, R.h., Bleichrodt, N., Resing, W.C.M., Lutje-Spelberg, H., Bruijn, E., de, and Zaal, J.N. (1984). *Vergelijkbaarheid van de WISC-R en RAKIT*. Research Instituut voor Onderwijs in het Noorden, Amsterdam: Vrije Universiteit.

Plomin, R., DeFries, J.C., and Loehlin, J.C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, 84, 309-322.

Plomin, R., and DeFries, J.C.(1983). The Colorado Adoption Project. *Child Development*, 54: 276-89.

Plomin, R., and DeFries, J.C.,(1985). Origins of Individual Differences in Infancy. In: Plomin, R. and DeFries, J.C. (Eds.). *Origins of Individual Differences in Infancy*, Orlando, pp. 48-64, 90-127.

Plomin, R. (1986). Multivariate analysis and developmental behavioral genetics: Developmental change as well as continuity. *Behavior Genetics*, 16, 25-43.

Plomin, R. (1986). *Development, genetics, and psychology*. Hillsdale: Erlbaum.

Plomin, R., and Daniels, D. (1987). why are children in the same family so different from one another? *Behavioral and Brain Sciences*, 10, 1-60.

Plomin, R., DeFries, J.C., and McClearn, G.E. (1990). *Behavioral Genetics: A Primer*. San Fransisco: Freeman.

Plomin, R., Campos, J., Corley, R., Emde, R.N., Fulker, D.W.,, Kagan, J., Reznick, J.S., Robinson, J., Zahn-Waxler, C. and DeFries, J.C. (1990). Individual differences during the second year of life: The MacArthur Longitudinal Twin Study. In: Columbo, J. and Fagan, J. (Eds.), *Individual differences in infancy: Reliability, stability, and predictability*. Hillsdale, New York.

Plomin, R., Emde, R.N., Braungart, J.M., Campos, J., Corley, R., Fulker, D.W., Kagan, J., Reznick, J.S., Robinson, J., Zahn-Waxler, C. (1993). Genetic change and continuity from fourteen to twenty months: the MacArthur Longitudinal Twin Study. *Child Development*, 64, 1354-1376.

Plomin, R., Reiss, D., Hetherington, E.M., and Howe, G.W. (1994). Nature and nurture:

Genetic contributions to measures of the family environment. *Developmental Psychology*, 30, 32-43. Plomin, R. (1995). Genetics and children's experiences in the family. *Journal of Child Psychology and Psychiatry*, 36, 33-68.

Plomin, R., DeFries, J.C., McClearn, G.E., Rutter, M. (1997). *Behavioral Genetics: a Primer*. Freeman, New York.

Plomin, R., Fulker, D.W., Corley, R. and DeFries, J.C.(1997). Nature, Nurture, and Cognitive Development from 1 to 16 Years: A Parent-Offspring Adoption Study. *Psychological Science*, 8, 442-447.

Plomin, R., DeFries, J.C., Mc Clearn, G.E., McGuffin, P., (2000). *Behavioral Genetics*, fourth edition, Worth Publishers, New York.

Plomin R, Hill L, Craig IW, McGuffin P, Purcell S, Sham P, Lubinski D, Thompson LA, Fisher PJ, Turic D, Owen MJ. (2001). A genome-wide scan of 1842 DNA markers for allelic associations with general cognitive ability: a five-stage design using DNA pooling and extreme selected groups. *Behavior Genetics*, 31(6), 497-509.

Plotsky, P.M., and Meany, M.J. (1993). Early, postnatal experience alters hypothalamic corticotrophin -releasing factor (CRF) mRNA, median eminence CRF content and stress-induced release in adult rats. *Mol. Brain Res.*, 18, 195-200.

Posener, J. A., DeBattista, C., Williams, G. H., Chmura Kraemer, H., Kalehzan, B. M., and Schatzberg, A. F. (2000). 24-Hour monitoring of cortisol and corticotrophin secretion in psychotic and nonpsychotic major depression. *Archives of General Psychiatry*, 57(8),755-760.

Posener, J. A., Schildkraut, J. J., Williams, G. H., and Schatzberg, A. F. (1997), Cortisol feedback effects on plasma corticotropin levels in healthy subjects. *Psychoneuroendocrinology*, 22(3),169-176.

Posthuma, D., van Baal, G.C.M., de Geus, E.J.C., Boomsma, D.I. (1999a). Association between APOE genotype and cognitive functioning in adult twins (abstract). *Psychophysiology*, 36, S91.

Posthuma, D., van Baal, G.C.M., de Geus, E.J.C., Snieder, H., Boomsma, D.I. (1999b). Association between apoE genotype, apoE-levels and cognitive performance. (abstract)., *Behavior Genetics*, 29, 366.

Posthuma, D., de Geus, E.J.C., Bleichrodt, N., and Boomsma, D. I. (2000). Twin-singleton differences in intelligence?, *Twin Research*, 3(2), 83-87.

Posthuma, D., and Boomsma, D. I. (2000). A note on the statistical power in extended twin designs. *Behavior Genetics*, 30(2),147-58.

Posthuma, D., Geus, E.J.C. de, Boomsma, D.I., (2001). Perceptual speed and IQ are associated trough common genetic factors. *Behavior Genetics*, 31(6), 593-602.

Posthuma, D. (2002). *Genetic variation and cognitive ability*. Thesis: Vrije Universiteit,

Amsterdam. The Netherlands.

Pritchard, J., Despres, J.P., Gagnon, J., Tchernof, A., Nadeau, A., Tremblay, A., Bouchard, C. 1998. Plasma adrenal, gonadal, and conjugated steroids before and after long-term overfeeding in identical twins. *J. Clin. Endocrinol. Metab.* 83, 3277-3284.

Pritchard, J., Despres, J.P., Gagnon, J., Tchernof, A., Nadeau, A., Tremblay, A., Bouchard, C., 1999. Plasma adrenal, gonadal, and conjugated steroids following long-term exercise-induced negative energy balance in identical twins. *Metabolism* 48, 1120-1127.

Raczynska, B., Lukaszewska, J., and Sklad, M. (1978). Plasma cortisol level in monozygotic and dizygotic twins. *Endokrynol Pol*, 29(1), 7-10.

Rende, R.D., Slomkowski, C.L., Stocker, C., Fulker, D.W., and Plomin, R. (1992). Genetic and environmental influences on maternal and sibling interaction in middle childhood: A sibling adoption study. *Developmental Psychology*, 28: 484-490.

Reul, J.M.H.M., and De Kloet, E.R. (1985). Two receptor systems for corticosterone in the rat brain: Microdistribution and differential occupation. *Endocrinology*, 117, 2505-2511.

Riad-Fahmy, D., Read, G.F., Walker, R.F., Griffiths, K., (1982). Steroids in saliva for assessing endocrine function. *Endocrine Review*, 3(4), 367-395.

Richman, N., Stevenson, J., and Graham, P.J. (1982). *Preschool to school: A behavioral study.* New York: Academic Press.

Rietveld, M.J.H., van der Valk, J.C., Bongers, I.L., Stroet, T.M., Slagboom, P.E., and Boomsma, D.I. (2000). Zygosity diagnosis in young twins by parental report. *Twin Research*, 3, 134-141.

Rietveld, M.J.H., Hudziak, J.J., Bartels, M., van Beijsterveldt, C.E.M., and Boomsma, D.I. (2002). Heritability of Attention Problems in Children: I. Cross-sectional results from a study of twins, age 3 to 12. Accepted, *Neuropsychiatric Genetics*.

Rietveld, M.J.H., Van Baal, G.C.M., Dolan, C.V., and Boomsma, D.I. (2000). Genetic Factor Analyses of Specific Cognitive Abilities in 5-Year-Old Dutch Children. *Behavior Genetics*, 30(1), 29-40.

Rietveld, M.J.H., Dolan, C.V., Van Baal, G.C.M., and Boomsma, D.I. (2002). A Twin Study of Differentiation of Cognitive Abilities in Childhood. *Behavior Genetics*, submitted.

Rijsdijk, F. (1997a). *The genetics of Neural Speed: A genetic study on nerve conduction velocity, reaction times, and psychometric abilites*. Thesis: Vrije Universiteit Amsterdam, The Netherlands

Rijsdijk, F. V., and Boomsma, D. I., (1997). Genetic mediation of the correlation between peripheral nerve conduction velocity and IQ. *Behavior Genetics*, 27(2),87-98.

Rijsdijk, F. V., Vernon, P. A., and Boomsma, D. I., (1998). The genetic basis of the

relation between speed-of-information-processing and IQ. *Behavioral Brain Research*, 95(1),77-84.

Rosmond, R., Chagnon, Y.C., Chagnon, M., Perusse, L., Bouchard, C., Bjorntorp, P. (2000[a]). A polymorphism of the 5'-flanking region of the glucocorticoid receptor gene locus is associated with basal cortisol secretion in men. *Metabolism*, 49, 1197-1199.

Rosmond, R., Chagnon, Y.C., Holm, G., Chagnon, M., Perusse, L., Lindell, K., Carlsson, B., Bouchard, C., Bjorntorp, P. (2000[b]). A glucocorticoid receptor gene marker is associated with abdominal obesity, leptin, and dysregulation of the hypothalamic-pituitary-adrenal axis. *Obes*, 8, 211-218.

Rosmond, R., Bjorntorp, P. (2000[c]). The hypothalamic-pituitary-adrenal axis activity as a predictor of cardiovascular disease, type 2 diabetes and stroke. *Journal Internal Medicine*, 246(2), 188-197.

Roy, M. P., Kirschbaum, C., and Steptoe, A. (2001). Psychological, cardiovascular, and metabolic correlates of individual differences in cortisol stress recovery in young men. *Psychoneuroendocrinology*, 26(4), 375-391.

Rowe, D.C. (1981). Environmental and genetic influences on dimensions of perceived parenting: A twin study. *Developmental Psychology*, 17, 203-208.

Rowe, D.C. (1983). A biometrical analysis of perceptions of family environment: A study of twin and singleton sibling kinship. *Child Development*, 54, 416-423.

Rowe, D.C., Woulbroun, E.J., and Gullwy, B.L. (1994). Peers and friends as nonshared environmental influences. In E.M. Hetherington, D. Reiss, and R. Plomin (Eds).,*Separate social worlds of siblings: The impact of nonshared environment on development* (pp. 159-173). Hillsdale, NJ: Erlbaum.

Rowe, D.C., and Kandel, D. (1997). In the eye of the beholder? Parental ratings of Externalizing and Internalizing Symptoms. *Journal of Abnormal Child Psychology*, 25, 265-275.

Rowe, D.C. (2001). *Biology and Crime*. Roxbury, Los Angeles.

Rubinow, D.R., Post, R.M., Savard, R., Gold, P.W. (1984). Cortisol hypersecretion and cognitive impairment in depression. *Archives of General Psychiatry*, 41(3), 279-83.

Ruiz, M., Lind, U., Gåfvels, M., Eggertsen, G., Carlstedt-Duke, J., Nilsson, L., Holtman, M., Stierna, P., Wikström A., Werner, S. (2001). Characterization of two novel mutations in the glucocorticoid receptor gene in patients with primary cortisol resitance. *Clinical Endocrinology*, 55(3), 363-371.

Rutter, M., Tizard, J., and Whitmore, K. (1970). *Education, Health* and *Behaviour*. Longman: London (reprinted 1981, Melbourne Krieger).

Rutter, M. (1985). Family and school influence on behavioural development. *Journal of Child Psychology and Psychiatry*, 26, 349-368.

Rutter, M., Macdonald, H., Le Couteur, A., Harrington, R., Bolton, P., and Bailey, A. (1990). Genetic factors in child psychiatric disorders-II. Empirical findings. *Journal of Child Psychology and Psychiatry*, 31, 39-83.

Scerbo, A.S., and Kolko, D.J. (1994). Salivary testosterone and cortisol in disruptive children relationship to aggressive, hyperactive, and internalizing behaviors. *Journal of the American Academy of Child and Adolescent Psychiatry*, 33(8), 1174-1184.

Scheer, F.A.J.L., Buijs, R.M. (1999). Ligth Affects Morning Salivary Cortisol in Humans. *Journal of Clinical Endocrinoly and Metabolism*, 84 (9), 3395-3398.

Schmidt, H., Schmidt, R., Fazekas, F., Semmler, J., Kapeller, P., Reinhart, B., and Kostner, G.M. (1996). Apolipoprotein E e4 allele in the normal elderly: Neuropsychologic and brain MRI correlates. *Clinical Genetics*, 50, 293-299.

Schmitz S., Fulker, D.W., and Mrazek, D.A., (1995). Problem behavior in early and middle childhood: An initial behavior genetic analysis. *Journal of Child Psychology and Psychiatry*, 36, 1443-1458.

Schmitz, S., Cherny, S.S., and Fulker, D.W. (1998). Increase in power through multivariate analyses. *Behavior Genetics*, 28, 357-364.

Schneider, B.H. (1993). *Children's Social Competence in Context: The Contribution of Family, School and Culture.* Oxford: Pergamon Press.

Schuckit, M.A., Hauger, R.L., Monteiro, M.G., Irwin, M., Duthie, L.A., Mahler, H.I. (1991). Response of three hormones to diazepam challenge in sons of alcoholics and controls. *Alcoholism, Clinical and. Experimental Research*, 15, 537-542.

Schulz, K. P., Halperin, J. M., Newcorn, J. H., Sharma, V., Gabriel, S. (1997). Plasma cortisol and aggression in boys with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(5), 605-609.

Schwartz, J. T., Reuling, F. H., Feinleib, M., Garrison, R. J., and Collie, D. J. (1972). Twin heritability study of the effect of corticosteroids on intraocular pressure. *Journal of Medical Genetics*, 9(2), 137-143.

Shih, J.C., Thompson, R.F. (1999). Monoamine oxidase in neuropsychiatry and behavior. *American Journal of Human Genetics*, 65(3), 593-8.

Sikich, L., and Todd, R.D. (1988). Are there neurodevelopmental effects of gonadal hormones related to sex differences in psychiatric ilness? *Psychiatric Development*, 4, 277-309.

Silberg, J.L., Erickson, M.T., Meyer, J.M., Eaves, L.J., Rutter, M.L., and Hewitt, J.K. (1994). The application of structural equation modeling to maternal ratings of twins' behavioral and emotional problems. *Journal of Consulting and Clinical Psychology*, 62, 510-521.

Silva, A.J., Frankland, P.W., Marowitz, Z., Friedman, E., Laszlo, G.S., Cioffi, D., Jacks, T.,

Bourtchuladze, R., Lazlo, G. (1997). A mouse model for the learning and memory deficits associated with neurofibromatosis type I. *Nature Genetics*, 15(3), 281-284.

Simonoff, E., Pickles, A., Hewitt, J., Silberg, J., Rutter, M., Loeber, R., Meyer, J., Neale, M., and Eaves, L. (1995). Multiple raters of disruptive child behavior: using a genetic strategy to examine shared views and bias. *Behavior Genetics*, 25, 311-326.

Simonoff, E., Pickles, A., Hervas, A., Silberg, J.L., Rutter, M., and Eaves, L. (1998). Genetic influences on childhood hyperactivity: contrast effects imply parental rating bias, not sibling interaction. *Psychological Medicine* **28**, 825-837.

Skodak, M. and Skeels, H.M.(1949). A Final Follow-up of One Hundred Adopted Children. *The Journal of Genetic Psychology*, 75, 85-125.

Smalley, S.L., Bailey, J.N., Palmer, C.G., Cantwell, D.P., McGough, J.J., Del'Homme, M.A., Asarnow, J.R., Woodward, J.A., Ramsey, C., Nelson, S.F. (1998). Evidence that the dopamine D4 receptor is a susceptibility gene in attention deficit hyperactivity disorder. *Molecular Psychiatry*, 3(5), 427-30.

Smyth, J. M., Ockenfels, M. C., Gorin, A. A., Catley, D., Porter, L. S., Kirschbaum, C., Hellhammer, D. H., and Stone, A. A. (1997). Individual differences in the diurnal cycle of cortisol. *Psychoneuroendocrinology*, 22(2), 89-105.

Solanto, M.V. (1998). Neuropsychopharmacological mechanisms of stimulant drug action in attention-deficit hyperactivity disorder: a review and integration. *Behavioral Brain Research*, 94(1):127-52

Spith-Schwalbe, E., Uthgenannt, D., Voget, D., Kern, W., Born, J., Fehm, H.-L. (1993). Corticotropin-releasing hormone-induced adrenocorticotropin and cortisol secretion depends on sleep and wakefulness. *Journal of Clinical Endocrinoly and Metabolism*, 77, 1170-1173.

Squire, L.R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological Reviews*, 99(2), 195-231.

Stallings, M.C., Cherny, S.S., Young, S.E., Miles, D.R., Hewitt, J.K., and Fulker, D.W. (1997). The familial aggregation of depressive symptoms, antisocial behavior, and alcohol abuse. *American Journal of Medical Genetics (Neuropsychiatric Genetics)*, 74, 183-191.

Starkman, M.N., Schteingart, D.E., Schork, M.A. (1981). Depressed mood and other psychiatric manifestations of Cushing's syndrome: relationship to hormone levels. *Psychosomatic Medicine*, 43(1), 3-18.

Stone, A.A., Schwartz, J.E., Smyth, J., Kirschbaum, C., Cohen, S., Hellhammer, D., Grossman, S. (2001). Individual differences in the diurnal cycle of salivary free cortisol: a replication of flattened cycles for some individuals. *Psychoneuroendocrinology*, 26(3), 295-306.

Stott, D.H. (1973). Follow-up study from birth of the effects of prenatal stresses. *Dev.*

Med. Child. Neurol, 15(6), 770-87.

Sunohara, G.A., Roberts, W., Malone, M., Schachar, R.J., Tannock, R., Basile, V.S., Wigal, T., Wigal, S.B., Schuck, S., Moriarty, J., Swanson, J.M., Kennedy, J.L., Barr, C.L. (2000). Linkage of the dopamine D4 receptor gene and attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39(12), 1537-1542.

Swanson, J.M., Sunohara, G.A., Kennedy, J.L., Regino, R., Fineberg, E., Wigal, T., Lerner, M., Williams, L., LaHoste, G.J., Wigal, S. (1998). Association of the dopamine receptor D4 (DRD4) gene with a refined phenotype of attention deficit hyperactivity disorder (ADHD): a family-based approach. *Molecular Psychiatryy*, 3(1), 38-41.

Tahir, E., Yazgan, Y., Cirakoglu, B., Ozbay, F., Waldman, I., Asherson, P.J. (2000). Association and linkage of DRD4 and DRD5 with attention deficit hyperactivity disorder (ADHD) in a sample of Turkish children. *Molecular Psychiatry*, 5(4), 396-404.

Tambs, K., Sundet, J.M., Magnus, P., and Berg, K. (1989). Genetic and environmental contributions to the covariance between occupational status, educational attainment, and IQ: A study of twins. *Behavior Genetics*, 19(2), 209-211.

Teasdale, T.W.,and Owen, D.R. (1984). Heredity and familial environment in intelligence and educational level- a sibling study. *Nature* 309: 602-622.

Tennes, K., Kreye, M., Avitable, N., Wells, R. (1986). Behavioral correlates of excreted catecholamines and cortisol in secondgrade children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 25, 764-770.

Tennes, K., and Kreye, M. (1985). Children's adrenocortical responses to classroom activities and tests in elementary school. *Psychosomatic Medicine*, 47, 451-460.

Thompson, L.A., Detterman, D.K., & Plomin, R. (1991). Association between cognitive abilities and scholastic achievement: Genetic overlap but environmental differences. *Psychological Science*, 2(3), 158-165.

Thompson, L.A. (1993). Genetic contributions to intellectual Development in infancy and Childhood. In: Vernon, P.A. (Ed.), *Biological Approaches to the Study of Human Intelligence*, pp. 95-138.

Trainer, P.J., McHardy, K.C., Harvey, R.D., Reid, I.W. (1993). Urinary free cortisol in the assessment of hydrocortisone replacement therapy. *Hormone and Metabolic Research*, 25, 117-120.

Trautman, P.D., Meyer-Bahlburg, H.F.L., Postelnek, J., and New, M.I. (1995). Effects of early prenatal dexamethasone on the cognitve and behavioral development of young children: Results of a pilot study. *Psychoneuroendocrinology*, 20, 439-449.

Tyrer, P. (1985). Neurosis divisible? *Lancet*,23(1) (8430), 685-688.

Vamvakpoulos, N.C., and Chrousos, G.P. (1994). Hormonal regulation of human

corticotropin-releasing hormone gene expression: implications for the stress response and immune/inflammatory reaction. *Endr. Rev.* **15**: 409-420.

Van Cauter, E. (1990). Diurnal and ultradian rhythms in human endocrine function: a minireview. *Hormone Reearch*, 34(2), 45-53.

Van Baal, G.M.C. (1997). *A genetic perspective on the developing brain*. Thesis: Vrije Universiteit Amsterdam, The Netherlands.

Van Beijsterveldt, C.E.M. (1996). *The genetics of electrophysiological indices of brain activity*. Thesis: Universiteit van Amsterdam, The Netherlands.

Van den Oord, E.J.C.G., Boomsma, D.I., Verhulst, F.C. (1994). A study of problem behaviors in 10- to 15-year-old biologically related and unrelated international adoptees. *Behavior Genetics*, 24(3): 349-357.

Van den Oord, E.J.C.G., Koot, H.M., Boomsma, D.I., Verhulst, F.C., and Orlebeke, J.F. (1995). A twin-singleton comparison of problem behaviour in 2-3-year-olds. *Journal of Child Psychology and Psychiatry*, 36: 449-458.

Van den Oord, E.J.C.G., Verhulst, F.C., and Boomsma, D.I. (1996). A genetic study of maternal and paternal ratings of problem behaviors in 3-year-old twins. *Journal of Abnormal Psychology*, 105, 349-357.

Van den Oord E.J.C.G., and Rowe D.C. (1997). Continuity and change in children's social maladjustment: a developmental behavior genetic study. *Developmental Psychology*, 33, 319-332.

Van den Oord, E.J.C.G., and Rowe, D.C. (1999). A cousin study of associations between family demographic characteristics and children's intellectual ability. *Intelligence* **27**, 251-266.

Van den Oord, E.J.C.G., Simonoff, E., Eaves, L.J., Pickles, A., Silberg, J., and Maes, H. (2000). An Evaluation of Different Approaches for Behavior Genetic Analyses with Psychiatric Symptom Scores. *Behavior Genetics*, 30(1), 1-18.

Van den Oord, E. J. C. G., Pickles, A. P., and Waldman, I. D. (2002). Normal variation and abnormality: An empirical study of the liability distributions underlying depression and delinquency. *Journal of Child Psychology and Psychiatry,* In press.

Van der Valk, J.C., Verhulst, F.C., Stroet, T.M., and Boomsma, D.I. (1998a). Quantitative genetic analysis of Internalizing and Externalizing Problems in a large sample of 3-year-old twins. *Twin Research*, 1, 25-33.

Van der Valk, J.C., Verhulst, F.C., Neale, M.C., and Boomsma, D.I. (1998b). Longitudinal genetic analysis of problem behaviors in biologically related and unrelated adoptees. *Behavior Genetics,* 28, 365-380.

Van der Valk, J.C., Verhulst, F.C., and Boomsma, D.I. (1999). Studying the development

of children's problem behaviors using quantitative genetic techniques. In: H.M. Koot, A.A.M. Crijnen, and R.F. Ferdinand (Eds.), *Child psychiatric epidemiology: Accomplishments* and *future directions* (pp. 116-141). Assen: Van Gorcum.

Van der Valk, J.C., van den Oord, E.J.C.G., Verhulst, F.C., and Boomsma, D.I. (2001) Using Parental Ratings to Study the Etiology of 3-year-old Twins' Problem Behaviors: Different Views or Rater Bias? *Journal of Child Psychology and Psychiatry*, 42(7), 921-931.

Van der Valk, J.C., van den Oord, E.J.C.G., Verhulst, F.C., and Boomsma, D.I. (2002, in press) Using common and unique parental views to study the etiology of 7-year-old twins' Internalizing and Externalizing problems.

Van der Valk, J.C., Van den Oord, E.J.C.G., Verhulst, F.C., and Boomsma, D.I. (2002). Genetic and environmental contributions to continuity and change of Internalizing and Externalizing Problems during childhood. *Journal of the American Acadamey of Child and Adolescent* Psychiatry, submitted.

Van Goozen, S.H.M., Matthys, W., Cohen-Kettenis, P.T., Gispen-de Wied, C., Wiegant, V.M., Engeland, H. (1998). Salivary Cortisol and Cardiovascular Activity during Stress in Oppositional-Defiant Disorder Boys and Normal Controls. *Biological Psychiatry*, 43, 531-539.

Van Haasen, P.P., De Bruyn, E.E.J., Pijl, Y.J., Poortinga, Y.H., Lutje-Spelberg, H.C., Vander Steene, G., Coetsier, P., Spoelders-Claes, R., Stinissen, J. (1986). *Wechsler Intelligence Scale for Children-Revised, Dutch Version*. Swets and Zetlinger B.V., Lisse.

Vanyukov, M.M., Moss, H.B., Plail, J.A., Blackson, T., Mezzich, A.C., Tarter, R.E. (1993). Antisocial symptoms in preadolescent boys and their parents: associations with cortisol. *Psychiatry Research*, 46, 9-17.

Veldhuis, H.D., Van Koppen, C., Van Ittersum, M., De Kloet, E.R. (1982). Specificity of the adrenal steroid receptor system in the rat hippocampus. *Endocrinology*, 110, 2044-2049.

Verhulst, F.C., Akkerhuis, G.W., and Althaus, M. (1985a). Mental health in Dutch children I. A cross cultural comparison. *Acta Psychiatrica Scandimavica*, 72, Supplement 323.

Verhulst, F.C., Berden, G., and Sanders-Woudstra, J.A.R. (1985b). Mental health in Dutch children II. Prevalence of psychiatric disorder and relationship between measures. *Acta Psychiatrica Scandimavica*, 72, Supplement 322.

Verhulst, F.C., and Akkerhuis, G.W. (1986). Mental health in Dutch children III. Behavioral -emotional problems reported by teachers of children aged 4-12. . *Acta Psychiatrica Scandimavica*, 72, Supplement 330.

Verhulst, F.C., Prince, J., Vervuurt-Poot, C., and de Jong, J. (1989). Mental Health in Dutch Adolescents: Self reported problems for ages 11-18. *Acta Psychiatrica Scandimavica*, 80, Supplement 356.

Verhulst, F.C., and Van der Ende, J. (1992a). Six-year developmental course of internalizing and externalizing problem behavior. *Journal of the American Academy of Child and Adolescent Psychiatry*, 31, 927-931.

Verhulst, F.C., and Van der Ende, J. (1992b). Six-year stability of parent reported problem behavior in an epidemiological sample. *Journal of Abnormal Child Psychology*, 20, 595-610.

Verhulst, F.C., van der Ende, J. (1993). "Comorbidity" in an epidemiological sample: a longitudinal perspective. *Journal of Child Psychology and Psychiatry*, 34(5), 767-783.

Verhulst, F.C., Koot, H.M., and Van der Ende, J. (1994). Differential predictive value of parents' and teachers' reports of children's problem behaviors: A longitudinal study. *Journal of Abnormal Child Psychology*, 22(5), 531-546.

Verhulst, F.C., and Van der Ende, J. (1995). The eight-year stability of problem behavior in an epidemiological sample. *Pediatric Research*, 38: 612-617.

Verhulst, F.C., Van der Ende, and Koot, H.M. (1996). *Handleiding voor de CBCL/4-18*. Afdeling Kinder- en Jeugdpsychiatrie, Sophia Kinderziekenhuis/ Academisch Ziekenhuis Rotterdam/ Erasmus Universiteit Rotterdam.

Verhulst, F.C., Van der Ende, J., Ferdinand, R.F., and Kasius, M.C. (1997). The prevalence of DSM-III-R diagnosis in a national sample of Dutch adolescents. *Archives of General Psychiatry*, 54: 329-336.

Verhulst, F.C. (1999). Fifteen years of research in chil psychiatric epidemiology. In: Koot, H.M., Crijnen, A.A.M., and Fredinand, R.F (eds). *Child Psychiatric Epidemiology: Accomplishments and future directions*. Van Gorcum, Assen , The Netherlands. p 16-38.

Vogler, G.P. and Fulker, D.W. (1983). Familial resemblance for educational attainment. *Behavior Genetics*, 13, 341-354.

Wadsworth, S.J., DeFries, J.C., Fulker, D.W. and Plomin, R. (1995a). Covariation among measures of cognitive ability and academic achievement in the Colorado Adoption Project: Sibling Analysis. *Personality and Individual Differences*, 18(1), 63-73.

Wadsworth, S.J., DeFries, J.C., Fulker, D.W. and Plomin, R. (1995b). Covariation among measures of cognitive ability and academic achievement in the Colorado Adoption Project: A Multivariate Genetic Analysis of Parent-Offspring and Sibling Data. *Behavior Genetics*, 25(1): 1-15.

Walker, B.R.. (1996). Abnormal glucocorticoid activity in subjects with risk for cardiovascular disease. *Endocrine Research*, 22(4), 701-708.

Walker, B.R., Phillips, D.I., Noon, J.P., Panarelli, M., Andrew, R., Edwards, H.V., Holton, D.W., Seckl, J.R., Webb, D.J., Watt, G.C. (1998). Increased glucocorticoid activity in men with cardiovascular risk factors. *Hypertension*, 31, 891-895.

Weitzman, E.D., Fukushima, D., Nogeire, C., Roffwarg, H., Gallagher, T.F., Hellman, L.,

(1971). Twenty-four Hour Pattern of the Episodic Secretion of Cortisol in Normal Subjects. *Journal of Clinincal Endocrinoly and Metabolism*, 33, 14-22.

Welberg, L.A. and Seckle. J.R. (2001). Prenetal stress, glucocorticoids and the programming of the brain. *Journal of Neuroendocrinology*, 13(2), 113-128.

Whelan, T.B., Schteingart, D.E., Starkman, M.N., Smith, A. (1980). Neuropsychological deficits in Cushing's syndrome. *J. Nerv. Ment. Dis.*, 168(12), 753-7.

Willerman, L., Horn, J.M., and Loehlin, J.C. (1977). The aptitude-achievement test distinction: A study of unrelated children reared together. *Behavior Genetics*, 7, 465-470.

Wilson, R.S. (1983). The Lopuisville Twin Study: developmental synchronies in behavior. *Child Development*, 54(2), 298-316.

Wolff, C.T., Friedman, S.B., Hofer, M.A., Mason, J.W. (1964). Relationship between psychological defenses and mean urinary 17-hydroxycorticosteroid excertion rates. *Psychosomatic Medicine*, 5, 576-591.

Wüst, S., Federenko, I., Hellhammer, D.H., Kirschbaum, C., (2000). Genetic factors, perceived chronic stress, and the free cortisol response to awakening. *Psychoneuroendocrinology*, 25, 707-720.

Yehuda, R., Bierer, L.M., Schmeidler, J., Aferiat, D.H., Breslau, I., Dolan, S. (2000). Low cortisol and risk for PTSD in adult offspring of holocaust survivors. *American Journal of Psychiatry*, 157, 1252-1259.

Yehuda, R., Kahana, B., Schmeidler, J., Southwick, S.M., Wilson, S., Giller, E.L. (1995). Impact of cumulative lifetime trauma and recent stress on current posttraumatic stress disorder symptoms in holocaust survivors. *American Journal of Psychiatry*, 152, 1815-1818.

Yehuda, R., Giller, E.L., Southwick, S.M., Lowy, M.T., Mason, J.W. (1991). Hypothalamic-pituitary-adrenal dysfunction in posttraumatic stress disorder. *Biological Psychiatry*, 30, 1031-1048.

Young, E. A., Lopez, J. F., Murphy-Weinberg, V., Watson, S. J., and Akil, H., (1998). The role of mineralocorticoid receptors in hypothalamic-pituitary-adrenal axis regulation in humans, *Journal of Clinical Endocrinology and Metabolism*, 83(9), 3339-3345.

Young, E.A., Aggen, S.H., Prescott, C.A., Kendler, K.S. (2000). Similarity in saliva cortisol measures in monozygotic twins and the influence of past major depression. *Biological Psychiatry*, 48, 70-74.

Zahn-Waxler, C., Schmitz, S., Fulker, D., Robinson, J., and Emde, R. (1996). Behavior problems in 5-year-old monozygotic and dizygotic twins: Genetic and environmental influences, patterns of regulation, and internalization of control. *Development and Psychopathology*, 8, 103-122.

List of publications

Journal Articles

Bartels, M., Van den Berg, M., Sluyter, F., Boomsma, D.I., and De Geus, E.J.C. (2002). Heritability of cortisol: Review and simultaneous analysis of twin studies. *Psychoneuroendocrinology*, in Press.

Bartels, M., Rietveld, M.J.H., Van Baal, G.C.M., Boomsma, D.I. (2002) Genetic and Environmental Influences on the Development of Intelligence. *Behavior Genetics*, 32(4), 237-249.

Bartels, M., Rietveld, M.J.H., Van Baal, G.C.M., Boomsma, D.I. (2002) Heritability of Educational Achievement in 12-year-olds and the overlap with Cognitive Abilities. *Twin Research*, 5(6), 1-10.

Bartels, M, De Geus, E.J.C., Kirschbaum, C., Sluyter, F. and Boomsma, D.I. Heritability of Daytime Cortisol in Children. *Behavior Genetics*, under revision.

Bartels, M., Boomsma, D.I., Rietveld, M.J.H., van Beijsterveldt, C.E.M., Hudziak, J.J., and E.J.C.G., van den Oord. Disentangling Genetic, Environmental, and Rater Effects on Internalizing and Externalizing Problem Behavior in 10-year-old Twins. *Psychological Medicine*, under revision.

Bartels, M., Van den Oord, E.J.C.G., Hudziak, J.J., Rietveld, M.J.H., van Beijsterveldt, C.E.M., and Boomsma, D.I. Genetic and environmental mechanisms underlying stability and change in problem behaviors at ages 3, 7, 10, and 12. *Developmental Psychology*, submitted

Bartels, M., J.J., Hudziak, Boomsma, D.I., Rietveld, M.J.H., van Beijsterveldt, C.E.M., and E.J.C.G., van den Oord. When mom's and dad's do and don't agree: A study of parent ratings of Internalizing and Externalizing behavior in 12-year-old twins. *Journal of the American Academy of Child and Adolescent Psychiatry,* submitted.

Bartels, M., Boomsma, D.I., Hudziak, J.J., Van Beijsterveldt, C.E.M., Rietveld, M.J.H., and Van den Oord, E.J.C.G. (2002). A longitudinal twin model for multiple raters: Illustrating the use of genetically informative designs for studying psychological data. *Psychological Methods*, submitted.

Boomsma, D.I., van Beijsterveldt, C.E.M., Rietveld, M.J.H., Bartels, M., and van Baal, G.C.M. (2001). Genetics mediate relation of birth weigth to childhood IQ. *Britisch Medical Journal*, 323, 1425.

Boomsma, D.I., Vink, J.M., van Beijsterveldt, C.E.M., de Geus, E.J.C., Beem,. A.L., Mulder, E.J.C.M., Riese, H., Willemsen, A.H.M., Bartels, M., van den Berg, M., Derks, E.M., de Graaff, S.C., Kupper, H.M., Polderman, J.C., Posthuma, D.I., Rietveld, M.J.H., Stubbe, J.H., Knol, L.I., Stroet, Th., van Baal, G.C.M. (2002). Netherlands Twin Register. *Twin Research*, in press

Rietveld M.J.H., Hudziak, J.J., Bartels, M., van Beijsterveldt, C.E.M., Boomsma. D.I.,

Heritability of Attention Problems in Children: I. Cross-sectional Results from a Study of Twins, age 3 to 12. *Neuropsychiatric Genetics*, in press

Book Chapter

Posthuma, D., Bartels, M., Boomsma, D.I. (2002). De gedragsgenetische benadering van veroudering. In: H. Schroots (Ed.) *Veroudering*. van Gorcum Uitgeverij. (in Dutch)

Abstracts

Bartels, M., Posthuma, D. & De Geus, E.J.C. (2000) Intelligence and peak EEG alpha frequency. *Journal of Psychophysiology*, 14(4), 260.

Bartels, M., Rietveld, M.J.H., Van Baal, G.C.M., and Boomsma, D.I (2000). Longitudinal Genetic Analyses of Intelligence at 5,7,10 and 12 years of age. Abstract BGA 2000, *Behavior Genetics*, 30(5), 398.

Bartels, M., De Geus, E.J.C., Kirschbaum, C., Rietveld, M.J.H., and Boomsma, D.I (2001). Abstract BGA 2001. Daytime Basal Cortisol and Common Childhood Psychopathology. *Behavior Genetics*, 31(5), 446.

Bartels, M., Rietveld, M.J.H., Hudziak, J.J., Boomsma, D.I., and Van den Oord, E.J.C.G. (2002). Internalizing and externalizing behavior: A Longitudinal perspective. Abstract BGA 2002, *Behavior Genetics*, in press

Boomsma, D.I., Beijsterveldt, C.E.M. van, Rietveld, M.J.H., Bartels, M., Baal, G.C.M. van. (2001). (abstract). Genetics mediate relation of birth weight to childhood IQ. *British Medical Journal*, 323, 1427.

Hudziak, J.J., Bartels, M., Rietveld, M.J.H., Stroet, Th.M., Todd, R., Stanger, C., Heath, A., Boomsma, D.I. (abstract). Univariate analyses of AP, AGG, and A/D in consecutive cohorts of 7, 10, and 12 year-old twin pairs: Gender and developmental findings. *American Journal of Medical Genetics*. 105 (7): O44a, 2001.

Hudziak, J.J., Rietveld, M.J.H., Bartels, M., Todd, R., Stanger, C., Heath, A., Boomsma, D.I. (abstract). The genetic and phenotypic relations between childhood attention, aggression, and anxious/depressed behavior: Evidence for developmental difference. *American Journal of Medical Genetics*. 105 (7): P131, 2001

Rietveld, M.J.H., Van Baal, G.C.M., Bartels, M., and Boomsma, D.I (2001). (abstract BGA 2001). Continuous and Discontinuous Cognitive Development. *Behavior Genetics*, 31(5), 467.

Dankwoord

Toen ik net begonnen was met het onderzoeksproject, wist ik al bijna wat ik in mijn dankwoord zou willen zeggen. Twee jaar later wist ik het precies, maar nu het moment werkelijk daar is moet ik er nog eens goed over nadenken. Het begin is eenvoudig. Het hele onderzoek had nooit uitgevoerd kunnen worden zonder de medewerking van vele tweelingparen en hun ouders. Zoals al uit de titel van dit proefschrift blijkt, is het een longitudinaal project, wat in de praktijk betekent dat de tweelingen en hun ouders meerdere keren gevraagd is om hun medewerking te verlenen. Zonder hen geen data en zonder data geen proefschrift. Ik wil alle tweelingen en hun ouders dan ook hartelijk danken voor hun medewerking. Maar met alleen data kom je ook nog niet echt ver. Ik wil dan ook als tweede Dorret Boomsma, mijn promotor, bedanken. Ik heb 4 jaar lang met heel veel plezier aan mijn onderzoek en proefschrift gewerkt en dat is grotendeels te danken aan jouw altijd inspirerende aanwezigheid. De snelheid en precisie waarmee jij dingen doorleest is gewoonweg indrukwekkend. Het lijkt net of jij veel meer uren in de dag hebt dan ieder ander. Ik heb alle vertrouwen in de toekomst; er valt nog heel veel te onderzoeken! Van mijn promotor wil ik dan overstappen op mijn copromotor, Edwin van den Oord. Edwin, alhoewel je pas de laatste 6 maanden (!) echt wist van het bestaan van mijn onderzoek is het toch gelukt om een enorme bijdrage te leveren. Mijn verblijf in Richmond is door jouw gastvrijheid, enthousiasme, kennis en onze samenwerking geweldig geweest. Onze planning van week 1 is voor de volle 100% uitgevoerd! Naast deze 'officiële' begeleiders is daar ook nog de 'officieuze'. Eco, jij zorgt voor een bruisende afdeling. Ik hoop dat ik ooit net zoveel uit mijn hoofd weet en zoveel kennis voor handen heb als jij (al is dat voor mij wat minder nodig omdat ik gewoon even het artikeltje opzoek in een georganiseerde ladekast). I like to thank Jim Hudziak for his 'clinical view' and all the effort he put in reading my papers. Via deze weg wil ik ook prof.dr. Buitelaar, prof.dr. Cohen-Kettenis, prof.dr. Delemarre-van de Waal, prof.em.dr. Orlebeke en prof.dr. Terwel hartelijk danken voor de vaak kostbare tijd en moeite die ze gestoken hebben in het lezen en beoordelen van dit proefschrift. I also like to thank prof.dr. Kirschbaum for his willingness to invest his time, his knowledge and experience into my thesis.

In de loop van de relatief 'eenzame' weg van de wetenschap ben ik er achter gekomen hoe waardevol collegae zijn. Alhoewel iedereen bezig is met zijn eigen stukje werk heb je ook weer een zelfde doel voor ogen. Mijn verblijf in Richmond met 'tea' maar zonder 'borrel' heeft mij des te meer doen beseffen dat we toch een mooi setje wetenschappers bij elkaar zijn. Marjolein, 4 jaar een kamer delen is veel meer dan toevallig dezelfde sleutel op zak hebben en uit hetzelfde raam kunnen kijken. Ik ben er trots op dat je mijn paranimf wilt zijn. Zolang de inzet van onze weddenschappen dezelfde blijft, moeten er nog heel wat afgesloten worden! Maar ook zonder het winnen of verliezen van de weddenschappen

hoop ik je nog vaak te zien, wat de toekomst ook moge brengen! Danielle, bedankt voor de gezelligheid, maar ook voor de hulp bij vele Mx probleempjes. Als ik een dringende vraag had en niet zeker wist of het een 'domme' vraag zou zijn, was daar altijd Caroline. Ik ging steeds opnieuw weg met een antwoord of een nieuw idee en was er dan nog steeds niet achter of het nu werkelijke een 'domme' vraag was of niet. Thérèse (en de andere medewerkersters van het NTR), ik wil niet eens proberen een uitspraak doen over het aantal vragenlijsten en aanverwante zaken dat door jullie handen is gegaan. Jullie waarde voor de afdeling is onbeschrijfelijk. Tinca, ik ben blij dat je de afdeling bent komen versterken. Een beetje Annewil, bitterballen of Til op z'n tijd is lang niet gek. Natascha, zonder jouw kennis van zaken, gepaste charmes en Sloterdijk-verhalen waren we nooit de afdeling geweest die we nu zijn. Eric, bedankt voor de altijd goede tips omtrent eet-technische zaken en bovenal voor het doorspitten van mijn verleden in de vorm van foto's voor het creëren van het omslag! Frans, my scientific-career-coach, neem nu maar van mij aan, het geluk ligt in de toekomst en het brein is net een harde schijf (wissen is eenvoudig). Nu zijn er natuurlijk nog vele collegae over en zonder iedereen persoonlijk te noemen, zeg ik alleen 'Keep up the good spirit'

Maar er is meer dan wetenschap en werk en alhoewel dit een dankwoord is bij een wetenschappelijk werk, is er een aantal mensen die op verschillende wijze ertoe hebben bijgedragen dat het leven zo mooi is als het is. Chronologische volgorde lijkt me een goede manier om deze groep te bedanken. Natuurlijk kan ik dan alleen maar bij 'mijn begin' beginnen. Lieve paps en mams, jullie onvoorwaardelijke steun, interesse, warmte en gezelligheid hebben mij grotendeels gemaakt tot wie ik ben. Lieve Jan Willem, Marlotte en Anouk, allemaal op jullie eigen wijze spelen jullie een belangrijke rol in mijn leven. Ik ben er trots op en enorm gelukkig mee dat ik 'zusje van Bartels' ben. Set, 1 keer prunussen trekken op de hei was genoeg. Mijn eeuwige rots in de branding! Sjaan, Jimster (Loes), Vrolijkje, Jet, Justa, Ka, Guul, Reuf, Haak, en Alwine, ieder op andere wijze, maar allemaal geweldig! Edwin, Esther, en Lyon (en natuurlijk Charlie), wat een bizarre tijd. De wonderen zijn de wereld nog niet uit. Ik wens jullie alle geluk van de wereld. Lieve Lyon, ik hoop dat we snel weer in de gelegenheid zijn om samen bellen te blazen.

Er is nog een aantal andere factoren die het leven geweldig maken. 'Zoals daar zijn' de ontdekking dat toeval niet bestaat. Voor mij is dit bewezen door het feit dat Mr Bojangles wordt bezongen maar ook een standbeeld heeft in Richmond. Tot slot wil ik nog toevoegen dat alle woorden, die hier toepasselijk zijn, terug te vinden zijn in het liedje van Alanis Morissette met als titel 'Head over Feet'.

Meike, 8 November 2002