CHAPTER 5

GENETIC AND ENVIRONMENTAL INFLUENCES ON THE STABILITY OF WITHDRAWN BEHAVIOUR IN CHILDREN: A LONGITUDINAL, MULTI-INFORMANT TWIN STUDY

This chapter is under revision as:
CHAPTER 5

GENETIC AND ENVIRONMENTAL INFLUENCES ON CHILDHOOD WITHDRAWN BEHAVIOUR

ABSTRACT

This study examined the contribution of genetic and environmental influences on the stability of withdrawn behaviour in childhood using a longitudinal multiple rater twin design. Maternal and paternal ratings on the withdrawn subscale of the Child Behavior Checklist were obtained from 14889 families when the twins were 3, 7, 10 and 12 years old. Withdrawn behaviour showed considerable stability throughout childhood, with correlation coefficients ranging from about .30 for the 9-year time interval to .65 for shorter time intervals. Individual differences in withdrawn behaviour were found to be largely influenced by genetic effects at all four time points, in both boys (50-66%) and girls (40-61%). Shared environmental influences explained a small to modest proportion (2 – 23%) of the variance at all ages and were slightly more pronounced in girls. Nonshared environmental influences were of moderate importance to the variance at all ages in both boys (21-38%) and girls (24-41%). The stability of withdrawn behaviour was largely explained by genetic effects, accounting for 74% of stability in boys and 65% in girls. Shared environmental effects explained 8% (boys) and 18% (girls) of the behavioural stability. As the main part of the shared environmental influences on stability was common to both raters, these effects could not be due to rater bias. Nonshared environmental effects accounted for the remaining covariance over time. Genetic and shared environmental correlations across age of the reliable phenotype approached unity, indicating that the same genes and shared environment influence withdrawn behaviour throughout childhood.

Key words: genetics; twins; childhood; problem behaviour; longitudinal studies; heritability

INTRODUCTION

Children scoring high on withdrawn behavioural scales are characterised by shy, inhibited, introvert and withdrawn behaviour. Withdrawn behaviour (WB) correlates with symptoms of anxiety and depression (Verhulst et al., 1996), and WB in childhood has been shown to predict anxiety disorders and major depression in adolescence and adulthood (Goodwin et al., 2004). In a follow-up study spanning 14 years, Hofstra et al. (2000) found that parent reported WB was an important predictor for malfunctioning in adulthood. Withdrawn behaviour at the time of first measurement predicted both adult internalising and externalising problems 14 years later. Furthermore, inhibited 3-year-olds (children who are shy, fearful and easily upset) were more likely to meet diagnostic criteria for depression when they were 21 years old (Caspi et al., 1996). Children described as “shy” on multiple time points showed increased incidence of anxiety problems in adolescence (Prior et al., 2000). The evidence that childhood WB is a predictor for anxiety and depression later in life is further supported by laboratory studies of behavioural inhibition. Behavioural inhibition (characterised by shy, inhibited behaviour, and fear for novel situations) is present in about 10 to 15% of children (Kagan et al., 1988). Behaviourally inhibited children have higher rates of childhood anxiety disorders (Biederman et al., 2003; Rosenbaum et al., 1993) and are at increased risk of developing adolescent social phobia (Hayward et al., 1998).

The continuity of problem behaviours advocates research into the underlying mechanisms influencing stability of behavioural traits. An extensive line of research indicates that behavioural problems in childhood show considerable continuity. For example, in a study using parent reported problem behaviours it was found that 41% of the children classified as deviant at first assessment also showed behavioural problems in the clinical range 14 years later (Hofstra et al., 2000). Moreover, behavioural observations at age 3 appeared to be predictive of psychiatric disorders 18 years later (Caspi et al., 1996). Stability of problem behaviours is not confined to clinical groups but is also found in general population samples. In a large population sample of Dutch children, a correlation of .48 for problem behaviours across an 8-year period was found (Verhulst et al., 1996). In the last decades, a range of longitudinal studies have focused on childhood externalising problem behaviours in general (e.g. Bartels et al., 2004b; Fergusson, 1998; Haberstick et al., 2005; Van der Valk et al., 2003a) and more specific problem behaviours such as attention problems (e.g. Mannuzza et al., 2003; Rietveld et al., 2004), aggression (e.g. Alink et al., 2006; Campbell et al., 2006) or conduct disorder (e.g. Fergusson et al., 2005; Kim-Cohen et al., 2003). Likewise, substantial attention has been devoted to internalising behaviours in general (e.g. Bartels et al., 2004b; Haberstick et al., 2005; Van der Valk et al., 2003a) and more
narrowly defined problems such as anxiety disorder and depression (e.g. Fombonne et al., 2001; Tram & Cole, 2006). However, surprisingly little research has focused on withdrawn behavioural problems.

A powerful way of unravelling the genetic and environmental effects on individual differences in the development of behavioural problems is the study of genetically related individuals. Both cross-sectional and longitudinal studies using the classical twin design have been conducted to assess heritability estimates for broad band internalising and externalising problem behaviours (Bartels et al., 2004b; Van der Valk et al., 2003a) as well as for specific syndrome scales such as aggression (Haberstick et al., 2006; Van Beijsterveldt et al., 2003), obsessive compulsive disorder (Hudziak et al., 2004; Van Groothuizen et al., 2007), juvenile bipolar disorder (Boomsma et al., 2006b), attention problems (Rietveld et al., 2004), and anxious/depressed (Boomsma et al., 2005). However, no large scale longitudinal twin studies into WB have been reported.

Family studies into childhood WB have been scarce, but there are indications that familial factors play a role. Behavioural inhibition is more frequent in children whose parents have agoraphobia and panic disorder (Rosenbaum et al., 1988), and anxiety disorders are more frequent in the families of behaviourally inhibited children (Rosenbaum et al., 1991). Furthermore, a study in a large sample of 4-year-old twins reported a heritability of 76% for shyness/inhibition, as assessed with a 3-item questionnaire (Eley et al., 2003). A few twin studies examined the cross-sectional heritability of WB at various ages in childhood using the Child Behavior Checklist (CBCL; Achenbach, 1991; Achenbach, 1992). An early twin study in a relatively small sample of 2 to 3 year-old twins found no significant genetic effects on variance in WB (Schmitz et al., 1995). Contrary to these findings, Van den Oord et al. (1996) reported major genetic influences (74%) and no evidence for shared environmental influences on individual differences in WB in a sample of 1358 3-year-old twin pairs. Eight years later, Derks et al. (2004) analysed data on WB of more than 9000 3-year-old twin pairs, including the data used in Van den Oord’s study and found moderate heritability (about 60% in boys; 45% in girls) and significant shared environmental effects. Two early twin studies examined the heritability of WB in middle childhood (sample sizes 181 and 203 pairs) and reported significant genetic effects (Edelbrock et al., 1995; Schmitz et al., 1995). On the other hand, a twin study from Taiwan including 279 12- to 16-year-old twin pairs (Kuo et al., 2004) found no significant genetic influences and major effects of shared and nonshared environment. One study compared WB data of biological and non-biological adopted siblings and found modest genetic influences at first assessment (age between 10-15 years) but no significant genetic effects three years later (Van der Valk et al., 1998).

These family studies are all based on parental ratings of WB. Using teacher report data of WB in 5-year-old twins, Polderman et al. (2006) found moderate genetic (49%) and nonshared environmental (51%) effects. Only two longitudinal studies (Schmitz et al., 1995; Van der Valk et al., 1998) have examined the genetic influences on the stability of childhood WB, and both failed to find significant genetic contributions to stability. However, in both studies the power to detect such effects was very low, due to limited sample size (Schmitz et al., 1995) or the design of the study (Van der Valk et al., 1998).

To summarise, the results of studies into the heritability of childhood WB have yielded varying results. Large scale studies into WB at later ages in childhood are lacking. Moreover, little is known about the genetic and environmental mechanisms underlying stability in WB. The aims of the current study are twofold. Firstly, this project, which is a follow-up of the twin sample studied by Derks et al. (2004), aims to examine the aetiology of variation in WB at various time points across childhood. Secondly, using the longitudinal nature of this study, we aim to assess the genetic and environmental factors underlying the stability of childhood WB.

Ratings of both maternal and paternal reported WB were incorporated in the analyses. Several studies into childhood behavioural problems have shown that different informants can provide different information about children’s behaviour (Achenbach & Rescorla, 2000; Achenbach & Rescorla, 2001; Seifge-Krenke & Kollmar, 1998; Van der Ende & Verhulst, 2005; Verhulst et al., 1996). Achenbach and Rescorla (2000; 2001) reported correlations between maternal and paternal ratings of WB of .69 for preschool children and of .57 for school-aged children. These correlations were based on data from a combined clinical and general population sample. In a general population only sample using the Dutch CBCL, the correlation between parental ratings of WB was found to be between .48 and .79 (Verhulst et al., 1996). The less-than-perfect correlation between parental ratings implies rater disagreement. Various studies have explored the sources of parental disagreement in ratings of problem behaviour (Bartels et al., 2003; Bartels et al., 2004a; Derks et al., 2004; Van der Valk et al., 2003b) using different structural equation models. Generally, it was found that parental agreement and disagreement was best explained by a psychometric model. This model, developed by Hewitt et al. (1992), assumes that parents not only assess the exact same behaviour of a child, but also rate an informant specific aspect of the child’s behaviour. This unique perception of the child’s behaviour can arise if the child behaves differently towards the different raters (e.g. the child is more withdrawn when it is with its mother than when it spends time with its father), or if the raters observe the child in different situations (e.g. the mother observes the child more often in the home environment, whilst the father often observes the child interacting with other children in the playground). Apart from these “real” differences
in behaviour, the unique perception of the child’s behaviour may also be influenced by rater bias and unreliability. Rater bias may arise if parents hold on to different normative standards, have specific response styles, or tend to stereotype the child’s behaviour. Unreliability may be an important source of rater disagreement if raters cannot give an accurate description of the behaviour under study. This may be relevant to our analyses, as some studies suggest that parents may be relatively insensitive to the more covert internalising problems of children (Ollendick & King, 1994; Seiffge-Krenke & Kollmar, 1998).

In the present study, stability of WB was assessed in longitudinal CBCL data from a large sample of 3, 7, 10 and 12 year-old twin pairs. The sample included roughly equal numbers of boys and girls. Genetic and environmental effects on stability in childhood WB were examined for both sexes. As both mother and father ratings of the twin’s behaviour were incorporated in the analyses, we controlled for rater bias effects by distinguishing between variance that is shared between parents (i.e. perception of the child’s behaviour common to both raters) and variance that is specific to one rater and might include variance due to rater bias. We compare the means and variance for WB in twins with estimates from a Dutch community sample (Van den Oord et al., 1995; Verhulst et al., 1996).

Methods

Participants

All participants were contacted via the Netherlands Twin register (NTR), kept by the Department of Biological Psychology at the VU University in Amsterdam (Bartels et al., 2007; Boomsma et al., 2002; Boomsma et al., 2006a). From 1986 onwards, the NTR has recruited families with multiples a few weeks or months after birth. Currently 40-50% of all multiple births are registered at the NTR. For the present study, data from twins born in 1986 – 2001 were included. Parents of the twins were asked to fill out a questionnaire assessing the twin’s behaviour at age 3, 7, 10 and 12 years. The questionnaires were mailed within three months of the twin’s 3rd, 7th, 10th and 12th birthdays. Two to three months after this mailing, reminders were sent to the non-responders. If finances permitted, persistent non-responders were contacted by phone. This procedure yielded a response rate between 61% and 73% (Bartels et al., 2007). Non-responders also include twin families who moved to an unknown address. From the original sample, 281 families were excluded because either one or both of the children had a disease or handicap that interfered with daily functioning. The total sample consisted of 14889 twin families. Ratings from both parents were available for 8479 families when the twins were 3 years old, 6414 at age 7, 4133 at age 10, and 2900 at age 12. Complete data from both parents at all time points were available for 1160 families. Maternal ratings were available for 14735 families, of which 13095 participated at age 3, 8855 at age 7, 5863 at age 10, and 3958 at age 12. Maternal data at all ages were available for 2797 families. Paternal ratings were available for 11499 families, of which 8794 families participated when the twins were 3 years old, 6522 at age 7, 4237 at age 10, 2974 at age 12. Complete father data on all four time points were available for 1290 families. This study is part of an ongoing project; the children born in later birth cohorts have not reached the age of 7, 10 or 12 years yet, which explains the decreasing numbers of participating families at the later ages.

To examine effects of sample attrition, we compared WB scores at age 3 of families who continued to participate at all other time points (when the twins were 7, 10, and 12 years of age), with families who only participated twice, once, or zero times at the subsequent time points. For the father ratings, there were no mean differences between these groups, neither for boys nor for girls. For the mother ratings, a significant effect of attrition was observed for both boys and girls. Mothers who continued to participate reported lower WB scores when their twins were 3 years old than mothers who did not participate at one or more of the later measurement occasions. However, these effects were small (effect size r=.07 in both boys and girls).

Of all participating twin pairs, 2310 were monozygotic males (MZM), 2591 were dizygotic males (DZM), 2619 monozygotic females (MZF), 2339 dizygotic females (DZF), 2566 opposite sex twins with a male firstborn (DOSMF), and 2464 opposite sex twins with a female firstborn (DOSFM). For 1380 same sex twin pairs, zygosity was based on DNA polymorphisms (n=1039) or blood group (n=341; Van Dijk et al., 1996). For the remaining same sex twin pairs (n=8479), zygosity was determined by discriminant analysis, using longitudinally collected questionnaire items. This method has proven to be of sufficient reliability: Rietveld et al. (2000) reported that agreement between this method and zygosity determination by blood/DNA polymorphisms was 93%.

Measures

Mother and father ratings of WB problems were obtained from the withdrawn / depressed syndrome scale of the CBCL/2-3 at age 3. The CBCL/2-3 (Achenbach, 1992) has been translated and validated for the Dutch population (Koot et al., 1997). The withdrawn/depressed scale in the CBCL/2-3 consists of 10 items. The parents were asked to rate the behaviour of the child on a 3-point scale based on the occurrence of the behaviour in the past 2 months. They were asked to rate the behaviour as 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.
were estimated for boys and girls separately. Furthermore, twin correlations at each age were estimated to gain insight in the contribution of genes and environment on the phenotypic variance that both raters agree on.

**Genetic modelling**

Since monozygotic (MZ) twins are (nearly) genetically identical, while dizygotic (DZ) twins on average only share 50% of their segregating genes, genetic model fitting of twin data allows for separation of the observed phenotypic variance into its additive genetic (A), shared environmental (C) and nonshared environmental (E) components. To incorporate the WB ratings from both parents into one model, a psychometric model was used (see Figure 1). The psychometric model enables a distinction between the variance that is shared by both raters and is independent of rater bias and unreliability (also called common phenotypic variance or reliable trait variance) and the variance that is rater specific (i.e. variance in the child’s behaviour that is uniquely perceived by one of the parents, also called unique or rater specific phenotypic variance). In the psychometric twin model, both the common and the unique phenotypic variance are decomposed into genetic, shared environmental and nonshared environmental influences. Significant genetic effects on the rater specific variance indicate that these unique perceptions of the child’s behaviour are “real”, as error and unreliability do not cause systematic effects and cannot mimic genetic influences. Shared environmental effects on the unique phenotype may be confounded by rater bias, as possible influences of rater bias will act independently of the zygosity of the twins. Unique nonshared environmental influences may be confounded by measurement error or unreliability.

To examine the stability of WB throughout childhood, the psychometric model was extended to incorporate data on all four time points. To gain insight in what factors are important for the continuity of WB, a fully parameterised model (the Cholesky decomposition) was used. This model served as a reference to examine the significance of the different components. The deterioration of the model fit was evaluated after each component was dropped from the fully parameterised model using $\chi^2$ tests. Genetic modelling was performed using Mx (Neale et al., 2006). In order to utilise all available data, including information of incomplete longitudinal data or data of which one of the parental ratings is missing, analyses were performed on the raw data.

**Data analyses**

Descriptive statistics, correlations and cross- correlations for WB at age 3, 7, 10 and 12 years were estimated using the software package Mx (Neale et al., 2006). To assess stability of WB over time, phenotypic correlations between the time points were estimated for boys and girls separately. Furthermore, twin correlations at each age and cross-twin/cross-age correlations were estimated for each zygosity group separately. These correlations give a first impression of the contribution of genetic and environmental effects on the variance of WB at each age, and on the stability of WB over time. Within-person inter-parent correlations were inspected to examine parental agreement on WB. Moreover, the cross-rater cross-twin correlations (e.g. the correlation between the mother rating of the oldest of the twins with the father rating of the youngest of the twins) were estimated to gain insight in the contribution of genes and environment on the phenotypic variance that both raters agree on.

**FIGURE 1. Psychometric model for multiple raters.**

A = Additive genetic effects; C = Shared environmental effects; E = Nonshared environmental effects.

At age 7, 10 and 12 WB was assessed using ratings of the WB syndrome scale of the CBCL/4-18 (Achenbach, 1991; Verhulst et al., 1996). The syndrome scale withdrawn encompasses 9 items, partially overlapping with the items in the CBCL/2-3. This time the parents were asked to rate the behaviour of the child in the preceding 6 months, on a 3-point scale identical to the scale used in the CBCL/2-3. For all ages, if more than two items on the WB scale were missing, the data were regarded incomplete and excluded from the analyses.
RESULTS

Table 1 shows the means and standard deviations for maternal and paternal ratings of withdrawn behaviour at age 3, 7, 10 and 12 for boys and girls separately. Withdrawn behavioural problems are significantly higher in boys than in girls at age 3 (p<.001 for both raters), age 10 (mother rating p=.04; father rating p=.01) and age 12 years (mother rating p<.001, but not at age 7 years (mother rating p=.24; father rating p=.72). As the power to detect differences was very high in our study, these mean differences are only of practical importance at age 3. Furthermore, ratings of WB in our twin sample are similar to the scores in the community sample at age 3, 7, and 10 years. At age 12, the mean scores are higher in the community sample.

The phenotypic correlations are given in Table 2 for both the mother and the father ratings, separately for boys and girls. These correlations give an indication of the stability of WB over time. Phenotypic correlations are around .30 between age 3 and later ages, and increase to .44 -.65 between age 7, 10, and 12 years. This pattern is similar in both parental ratings and is observed in both boys and girls. Within-person cross-rater correlations (not shown in Table 2) were similar across zygosity and sex and were on average estimated as r=.53 at age 3, .56 at age 7, .57 at age 10 and .60 at age 12.

The twin correlations and cross-twin-cross-age correlations are presented in Table 3 for both the mother and the father ratings. Inspection of the MZ and DZ twin correlation (on the diagonal) at the four time points gives a first impression of what factors influence variance in WB. At all ages, MZ correlations are higher than DZ correlations in both sexes, indicating that genetic factors play a role. Twin correlations in opposite sex twins are similar to the correlations in DZ same sex twins, suggesting that there are no sex differences in genes or shared environment influencing WB. Apart from age 7, the MZ correlations are not twice as high as the DZ correlations, suggesting that shared environmental factors also play a role. Inspection of the MZ and DZ cross correlations (off-diagonal in Table 3) can provide insight in what fac-
tors are important for the stability of WB over time. As compared to the DZ cross correlations, MZ cross correlations are slightly higher between age 3 and subsequent ages, and considerably higher between age 7 and later ages, indicating genetic effects on stability. However, MZ cross correlations are not twice as high, particularly between age 3 and later ages, suggesting that shared environmental effects on stability are also important.

Table 4 displays the cross-twin-cross-rater correlations within age (diagonal) and across age (off-diagonal). These correlations yield a first impression on the importance of genes and environment on the common phenotypic variance, and thus the reliable trait variance. At all ages, the MZ cross-rater correlations are larger than the DZ cross-rater correlations, indicating genetic effects on the common phenotype. Apart from the correlations at age 7, the DZ correlations are higher than would be expected based on genetic influences alone, therefore shared environmental influences also seem to influence the common phenotypic variance. For all ages the cross-twin-cross-rater correlations are lower than the cross-twin-within-rater correlations given in Table 3. These differences indicate parental disagreement, and reveal the part of the total variance that is due to a specific rater. Similar to the pattern of the within age twin correlations, the cross-twin-cross-rater correlations across age are higher in MZ than in DZ twins, indicating genetic effects. These correlations are less than...
twice as high in MZ twins compared to DZ twins, especially in girls, indicating that
shared environmental influences also play a role in the stability of the common phe-
notype. The cross-rater-cross-age correlations are similar to the within-rater-cross-
age correlations between age 3 and later ages. This pattern indicates that most of the
stability of WB is perceived by both raters. In later phases of childhood, the within-
rater stability increases and becomes more pronounced, indicating that the common
phenotype show considerable con-
tinuity over time.

The significance of all genetic and environmental components was tested by ex-
amining the deterioration of the model fit after each component was dropped from
the fully parameterised model. All variance components were found to be significant,

| TABLE 5. Relative contributions of genetic (A), shared (C) and nonshared (E) envi-
ronmental influences to the total (common + unique) variances (diagonal) and cov-
ariances (off-diagonal) of withdrawn behaviour for boys (above diagonal) and girls
(below diagonal). |
<table>
<thead>
<tr>
<th>Age</th>
<th>3</th>
<th>7</th>
<th>10</th>
<th>12</th>
<th>3</th>
<th>7</th>
<th>10</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong></td>
<td>.61/.62</td>
<td>.83</td>
<td>.72</td>
<td>.75</td>
<td>.51/.61</td>
<td>.87</td>
<td>.72</td>
<td>.81</td>
</tr>
<tr>
<td>7</td>
<td>.71</td>
<td>.60/.64</td>
<td>.80</td>
<td>.68</td>
<td>.75</td>
<td>.62/.66</td>
<td>.78</td>
<td>.71</td>
</tr>
<tr>
<td>10</td>
<td>.64</td>
<td>.74</td>
<td>.58/.59</td>
<td>.68</td>
<td>.64</td>
<td>.63</td>
<td>.44/.52</td>
<td>.48</td>
</tr>
<tr>
<td>12</td>
<td>.55</td>
<td>.61</td>
<td>.67</td>
<td>.46/.53</td>
<td>.60</td>
<td>.67</td>
<td>.43</td>
<td>.40/.50</td>
</tr>
<tr>
<td><strong>C</strong></td>
<td>.15/.09</td>
<td>.05</td>
<td>.10</td>
<td>.13</td>
<td>.23/.18</td>
<td>.01</td>
<td>.10</td>
<td>.09</td>
</tr>
<tr>
<td>7</td>
<td>.20</td>
<td>.05/.02</td>
<td>.01</td>
<td>.07</td>
<td>.14</td>
<td>.05/.03</td>
<td>.04</td>
<td>.04</td>
</tr>
<tr>
<td>10</td>
<td>.21</td>
<td>.06</td>
<td>.05/.06</td>
<td>.07</td>
<td>.21</td>
<td>.17</td>
<td>.19/.11</td>
<td>.22</td>
</tr>
<tr>
<td>12</td>
<td>.33</td>
<td>.16</td>
<td>.12</td>
<td>.19/.10</td>
<td>.27</td>
<td>.07</td>
<td>.25</td>
<td>.19/15</td>
</tr>
<tr>
<td>7</td>
<td>.09</td>
<td>.35/.34</td>
<td>.19</td>
<td>.24</td>
<td>.11</td>
<td>.32/.31</td>
<td>.18</td>
<td>.25</td>
</tr>
<tr>
<td>10</td>
<td>.15</td>
<td>.20</td>
<td>.37/.36</td>
<td>.25</td>
<td>.15</td>
<td>.20</td>
<td>.37/37</td>
<td>.30</td>
</tr>
<tr>
<td>12</td>
<td>.12</td>
<td>.23</td>
<td>.21</td>
<td>.35/38</td>
<td>.13</td>
<td>.26</td>
<td>.32</td>
<td>.41/35</td>
</tr>
</tbody>
</table>

*First figure is the relative contribution for girls, second figure for boys.

| TABLE 6. Relative contributions of genetic (A), shared (C) and nonshared (E) envi-
ronmental influences to the variances (diagonal) and covariances (off-diagonal) of
withdrawn behaviour for the common phenotype and the unique (rater specific)
phenotype for boys (above diagonal) and girls (below diagonal). |
<table>
<thead>
<tr>
<th>Age</th>
<th>3</th>
<th>7</th>
<th>10</th>
<th>12</th>
<th>3</th>
<th>7</th>
<th>10</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong></td>
<td>.28/.33</td>
<td>.83</td>
<td>.72</td>
<td>.75</td>
<td>.30/.36</td>
<td>.87</td>
<td>.72</td>
<td>.77</td>
</tr>
<tr>
<td>7</td>
<td>.67</td>
<td>.35/.40</td>
<td>.46</td>
<td>.68</td>
<td>.75</td>
<td>.39/.42</td>
<td>.54</td>
<td>.68</td>
</tr>
<tr>
<td>10</td>
<td>.64</td>
<td>.35</td>
<td>.17/.21</td>
<td>.32</td>
<td>.64</td>
<td>.42</td>
<td>.20/25</td>
<td>.41</td>
</tr>
<tr>
<td>12</td>
<td>.55</td>
<td>.59</td>
<td>.24</td>
<td>.22/27</td>
<td>.60</td>
<td>.67</td>
<td>.30</td>
<td>.25/30</td>
</tr>
<tr>
<td><strong>C</strong></td>
<td>.09/.04</td>
<td>.01</td>
<td>.10</td>
<td>.07</td>
<td>.10/.05</td>
<td>.01</td>
<td>.10</td>
<td>.07</td>
</tr>
<tr>
<td>7</td>
<td>.09</td>
<td>.01/.00</td>
<td>.00</td>
<td>.00</td>
<td>.10</td>
<td>.01/.00</td>
<td>.01</td>
<td>.01</td>
</tr>
<tr>
<td>10</td>
<td>.21</td>
<td>.05</td>
<td>.05/.03</td>
<td>.03</td>
<td>.21</td>
<td>.05</td>
<td>.07/03</td>
<td>.04</td>
</tr>
<tr>
<td>12</td>
<td>.24</td>
<td>.06</td>
<td>.10</td>
<td>.09/02</td>
<td>.27</td>
<td>.06</td>
<td>.13</td>
<td>.10/02</td>
</tr>
<tr>
<td><strong>E</strong></td>
<td>.07/09</td>
<td>.11</td>
<td>.18</td>
<td>.10</td>
<td>.08/09</td>
<td>.12</td>
<td>.18</td>
<td>.10</td>
</tr>
<tr>
<td>7</td>
<td>.09</td>
<td>.12/13</td>
<td>.06</td>
<td>.24</td>
<td>.11</td>
<td>.13/14</td>
<td>.07</td>
<td>.25</td>
</tr>
<tr>
<td>10</td>
<td>.15</td>
<td>.04</td>
<td>.04/05</td>
<td>.04</td>
<td>.15</td>
<td>.05</td>
<td>.04/06</td>
<td>.05</td>
</tr>
<tr>
<td>12</td>
<td>.12</td>
<td>.23</td>
<td>.05</td>
<td>.10/10</td>
<td>.13</td>
<td>.26</td>
<td>.05</td>
<td>.11/11</td>
</tr>
</tbody>
</table>

Note: **A** = Additive genetic influences on the common phenotype; **C** = Shared envi-
ronmental influences on the common phenotype; **E** = Nonshared environ-
mental influences on the common phenotype; **A** = Additive genetic influences on
the unique phenotype; **C** = Shared environmental influences on the unique phenotype; **E** = Nonshared environmental influences on the unique phenotype. *First figure is the relative contribution for girls, second figure for boys.
for both the common phenotype and the unique phenotype (p<.001 for all components). Table 5 presents the relative contributions of genetic, shared and nonshared environmental influences to the total (common + unique) variances of WB at each age (diagonal) and to the stability of WB over time (off-diagonal). The heritability of both paternal and maternal rated behaviour is about 60% in both sexes at age 3 and age 7, and decreases slightly to 40-53% at the later ages. Shared environmental effects are of modest importance for the variance in both sexes, although the influence is slightly larger in girls. Nonshared environmental influences explain 21-41% of the variance at all ages. The relative importance of A, C and E to the total covariances between the four time points are shown on the off-diagonals of Table 5. Genetic effects are the driving force behind the stability of WB. Following the mother rat-

ings of WB in boys, on average 74% of the stability ((83% + 72% +75% + 80% + 68% + 68%/6 = 74%) is accounted for by genetic effects. Likewise, 65% of the maternal rated stability of WB in girls is explained by genetic effects. These contributions are similar for the paternal ratings. Shared environmental influences are important for the stability of WB in girls, and explain about 18% of the behavioural continuity in both mother and father ratings. In boys, these effects only explain about 8% of the covariance. The remaining part of the covariance is explained by nonshared environmental effects.

In Table 6, a distinction is made between the contributions of A, C and E to the common phenotype (A\textsubscript{c}, C\textsubscript{c}, and E\textsubscript{c},) and to the phenotype unique to either the mother or the father ratings (A\textsubscript{uf}, C\textsubscript{uf}, E\textsubscript{uf}). At age 3 and age 7, the common and the unique phenotypic variance contribute equally to the total variance. At later ages in childhood, substantial extra information is added by the specific raters, especially by the mothers. A large proportion of the rater specific variance is due to genetic influences, indicating that this rater specific information of the child’s behaviour is real. About half of the shared environmental influences on WB is rater specific. These effects may be real rater specific shared environmental effects, but may also be due to rater bias. The off-diagonals of Table 6 show the influences of common and rater specific genetic and environmental influences on the covariances. Common genetic influences are most important for explaining continuity of WB, and account for about 55% of the total covariance in girls and 65% in boys. Table 6 also shows that most of the shared environmental influences to stability are common to both raters. This indicates that these effects are “real” and not due to rater bias. On the whole, the possible effects of rater bias are small in this study, as rater specific shared environmental effects on average only explain 4% (boys) to 5% (girls) of the stability.

Lastly, the correlations between the genetic influences over time and the correlations between shared and nonshared environmental effects across development are displayed in Table 7. The genetic correlation of the common phenotype remains high across time, indicating that roughly the same genes influence the stability of the reliable WB phenotype between age 3 and age 12 years. Likewise, the shared environmental effects on the stability of the common phenotype remain largely the same over time. The nonshared environmental correlations and the rater specific genetic and shared environmental correlations over time are lower. These effects are thus more variable over time.
DISCUSSION

We studied the aetiology of WB in a large sample of 3-, 7-, 10-, and 12-year-old twins, and explored the genetic and environmental influences on stability of WB across childhood. Both maternal and paternal ratings of their children’s behaviour were analysed in order to identify the part of the phenotype that both raters agree on, and correct for possible rater bias.

Individual differences in WB were largely influenced by genetic effects at all ages, in both boys (heritability estimates 50-66%) and girls (heritability estimates 40-61%). Shared environmental influences explained a small to modest proportion (2-23%) of the variance at all ages in both sexes, but were slightly more pronounced in girls. Nonshared environmental influences were of moderate importance to the variance at all ages in both boys (21-38%) and girls (24-41%). This study is the first large scale study examining genetic and environmental influences on WB at multiple time points in childhood which is adequately powered to test for shared environmental effects. Results from previous family studies into childhood WB provided varying results. Two earlier twin studies in 3-year-olds from the NTR (Derks et al., 2004; Van den Oord et al., 1996), found significant genetic effects, varying in magnitude from 45 to 74%. The most recent study with the largest sample size (Derks et al., 2004) also found significant shared environmental effects that were more pronounced in girls than in boys. Two studies in middle to late childhood reported heritabilities of 40% (Schmitz et al., 1995) and 53% (Edelbrock et al., 1995). Two other family studies in middle to late childhood found no or little genetic effects (Kuo et al., 2004; Schmitz et al., 1995; Van der Valk et al., 1998). In the Taiwanese twin study (Kuo et al., 2004) the 95% confidence interval for additive genetic effects, however, was between 0 and 55%, indicating that the proportion of the variance explained by additive genetic influences may have been as large as 55%. On the other hand, the different results in the Taiwanese study compared to ours may also be due to cultural differences between the populations. Previous studies have suggested cultural effects on WB (Crijnen et al., 1999; Murad et al., 2003).

The longitudinal nature of the current study allowed for examination of the stability of WB throughout childhood. Withdrawn behaviour showed considerable continuity over time, with stability coefficients ranging from .30 between age 3 and 12 years to .65 for the shorter time intervals. These phenotypic correlations were similar to the correlations reported in other studies of childhood WB. In a small longitudinal twin sample, Schmitz et al. (1995) found a correlation of .33 between CBCL/2-3 scores and CBCL/4-18 scores of WB. In a large general population sample from the Netherlands (Verhulst et al., 1996), an 8-year stability coefficient of .36 was reported. Smaller time intervals gave higher stability coefficients (.47 for a 6-year interval; .46 for a 4 year interval and .60 for a 2-year interval).

We studied the genetic and environmental influences underlying stability in childhood WB. Stability in WB problems was largely accounted for by genetic effects, in boys (about 74% across time) and girls (about 65%). In girls, the shared environment was of moderate importance for continuity of WB, these influences explained on average 18% of the stability over time. In boys, these effects were less important, explaining about 8% of the stability. Interestingly, most of the shared environmental influences on stability were common to both raters, indicating that these effects are not due to rater bias. Nonshared environmental effects explained 17 to 20% of the stability over time in both sexes. Around two third of these effects were common to both raters.

In a longitudinal study into childhood internalising behaviour in this same sample, Bartels et al. (2004b) found that stability in internalising problems was accounted for by both genetic and shared environmental effects, and these effects were roughly of the same importance for stability (43 vs. 47%). We found that genetic effects largely explained stability in WB (65% in girls; 74% in boys), whilst shared environmental influences were only of modest importance (8-20%). The broad band internalising problem behaviour scale of the CBCL/2-3 includes the Withdrawn and Anxious/Depressed syndrome scales. The internalising problems scale in the CBCL/4-18 consists of the subscales Anxious/Depressed; Somatic complaints and Withdrawn. Bartels et al. (2004b) found decreasing genetic effects and increasing shared environmental effects on the variance in internalising behaviour over time. The same pattern was found in a longitudinal study into Anxious/Depressed behaviour in childhood (Boomsma et al., 2005). Our study suggest that, unlike the other syndrome scales that make up the broad band internalising scale, for WB shared environmental effects do not become increasingly important in later childhood.

The correlation between the genetic influences on the common phenotype over the course of development approaches unity. As the common phenotype represents the behaviour that both raters agree upon, and can thus be considered as a reliable phenotype, the high genetic correlations suggest that the same genes influence the continuity of WB over time. Similarly, the shared environmental correlation of the common phenotype is close to 1.0 over time, indicating that a stable persistent shared environmental influence is of importance for behavioural stability. Nonshared environmental correlations on the other hand vary over time in both boys and girls, suggesting that these effects are less persistent. The family environment could be an important factor for stability. Some studies suggest that parental behaviour may moderate the stability of behavioural inhibition and shyness in young children. Inappropriate affectionate parenting (Park et al., 1997) or maternal over-control
(Rubin et al., 2002) could increase stability of these behaviours. The family environment can show up in the shared environmental component (i.e. socioeconomic status of the family, parental rearing practices) or in the nonshared environmental component (i.e. child specific parenting). Nonshared environmental influences, such as traumatic experiences, the consequences of an accident or illness could also account for stability in WB.

The extent to which a child displays WB might be influenced by the composition of the family in which the child is raised. Our study focused on variation in WB in twins. It may be that twins show less WB compared to singletons, because they are raised with a sibling of the exact same age. On the other hand, if twins mainly interact with each other in childhood, twins may be more inhibited than non-twin siblings or singletons in interaction with others than their co-twin. Comparing the mean scores of our twin sample with the scores in a Dutch community sample showed that withdrawn scores are similar at age 3, 7 and 10 years. At age 12, however, mean problem scores were higher in the community sample than in the twin sample. In a large twin-singleton comparison study (Pulkkinen et al., 2003), 12-year-old twins were reported to be more socially adaptable than non-twins, but no twin-singleton differences were found for social anxiety. A twin-singleton comparison of both maternal CBCL withdrawn ratings and laboratory assessment of inhibition in 5-year-olds yielded inconsistent results (DiLalla & Caraway, 2004). According to laboratory ratings, twins were more inhibited than non-twins, whilst maternal ratings showed the opposite. These studies yield no explanation for the low mean withdrawn scores observed in our twin sample at age 12. To explore possible twin-singleton differences, future studies are needed.

The current study highlights the importance of genetic effects on stability in WB. Unlike the other internalising behaviours, shared environmental influences do not become increasingly important throughout childhood. Results from longitudinal (Bongers et al., 2003) and cross-sectional studies (Achenbach, 1991; Verhulst et al., 1996) indicate an increase in WB from childhood to adolescence. Future longitudinal research should extend the current study and investigate stability of WB into adolescence. Since childhood WB has been shown to be a predictor for anxiety disorders and depression later in life (Goodwin et al., 2004), insight into the developmental mechanisms underlying stability of WB into adolescence would be highly desirable. This is particularly true given the differences between the Anxious/Depressed and WB scales of the broad band internalising scale. It appears from our studies that WB has a different genetic architecture and longitudinal course than Anxious/Depressed. While Anxious/Depressed has received a great deal of research attention, there has been relatively little investigation into the long term sequela of WB. Our research provides evidence for the need for further work on this phenotype.

**References**


CHAPTER 5


Verhulst, F. C., Van der Ende, J., & Koot, H. M. (1996). Handleiding voor de CBCL/4-18 [Dutch manual for the CBCL/4-18]. Rotterdam, the Netherlands: Academic Medical Centre Rotterdam / Erasmus University, Sophia Children’s Hospital, Department of Child Psychiatry.