Chapter **7**

General discussion

In this thesis I investigated the development of internalizing problems, externalizing prob-lems, and ADHD symptoms during childhood, in twins and singletons. One assumption that is made when using twin data is that results from twin samples can be generalized to single-ton populations. In this thesis, the validity of this assumption was examined. An important life event for children as they grow up is parental divorce, and I investigated the influence of parental divorce on degrees of internalizing and externalizing problems at the ages of 3 and 12 years. Because life events such as parental divorce can also act as modifiers of the genetic architecture of complex traits, I also looked at the interaction of parental divorce and the influence of genetic factors on internalizing and externalizing problems.

The data for this thesis come from two large longitudinal studies: the Netherlands Twin Register (NTR; VU University, Amsterdam) and the Zuid-Holland Study (ZHS; Erasmus Medical Center, Rotterdam). The NTR supplied data on twins and the ZHS data on singletons. Based on these data, latent growth curves were estimated for internalizing and externalizing problems for twins and singletons (chapter 3). Growth mixture modeling was used to determine devel-opmental trajectories of ADHD symptoms for twins and singletons (chapter 4). The influence of parental divorce on levels of internalizing and externalizing problems was investigated with regression and survival analysis (chapter 5). Finally, the influence of parental divorce on the genetic architecture of internalizing and externalizing problems was examined (GxE, chapter 6). The results are discussed below, as well as the strengths and limitations of the studies, and implications and suggestions for future research.

Comparing twins to singletons

Internalizing problems

We studied the development of internalizing problems from ages 6 to 12 years in chapter 3. The main finding was that these problems develop similarly for twins and singletons up to age 9. After this age twins' levels of internalizing problems started to decrease, whereas in singletons they stabilized. This resulted in a lower level of internalizing problems for twins than singletons by the age of 12 years. In early adolescence, twinship may be a protective factor to the development of internalizing problems. This may be explained by the premise that twins have someone close for support. Siblings, in general, have been found to be a source of support to each other (Furman & Buhrmester, 1985). As twins appear to have a more intimate relationship with each other than regular siblings (Segal, Chavarria, & Hoven Stohs, 2008), they could be a greater source of support to each other than regular siblings. Higher levels of sibling support have been associated with lower levels of internalizing problems (Branje, Van Lieshout, Van Aken, & Haselager, 2004). Sibling support is present at younger ages as well, but may be stronger in early adolescence (Branje, et al., 2004). This may explain

why twins show less internalizing problems than singletons by the age of 12 years and not already at a younger age. Also, given that conflicts in sibling relationships predict internalizing problems (Moser & Jacob, 2002), one could hypothesize that twin-pairs may experience less conflicts in their relationships than other sibling pairs, which may explain the lower levels of internalizing problems in twins.

Externalizing problems

The development of externalizing problems from ages 6 to 12 years was similar for twins and singletons. Externalizing problems decreased steadily with age for both twins and singletons, and their developmental trajectories were not significantly different with respect to the initial levels of problems and the developmental course with age. Although decreasing levels of externalizing problems in childhood have consistently been found in earlier studies (e.g.,Bongers, Koot, Van der Ende, & Verhulst, 2003; Leve, Kim, & Pears, 2005; Miner & Clarke-Stewart, 2008; Stanger, Achenbach, & Verhulst, 1997), our findings are innovative as developmental trajectories for twins and singletons had not yet been compared. The main finding that externalizing problems develop similarly for twins and singletons up to age 12 is in agreement with earlier cross-sectional studies among 11- and 12-year-old Finnish twins and singletons (Pulkkinen, et al., 2003) and 5- to 15-year-old Norwegian twins and singletons (Gjone & Novik, 1995) that reported no evidence for twin-singleton differences in externalizing problems. Although other cross-sectional twin-singleton comparisons have shown some mixed results (i.e., lower or higher levels of externalizing problems in twins than in singletons (Moilanen, et al., 1999; Van den Oord, et al., 1995, Gau, et al., 1992), my longitudinal study contributes to a further confirmation of the generalizability of twins to singletons with respect to externalizing problems.

ADHD symptoms

Chapter 4 reports on three trajectories of mother-rated CBCL Attention Problems in boys and girls from 6 to 12 years: stable low, low-increasing and high-decreasing symptom levels. The Attention Problems scale of the CBCL covers both attention problems and hyperactivity symptoms, and will be further referred to as ADHD symptoms. The finding of these three trajectories is in agreement with the study from Malone and colleagues (2010) which identified three trajectories that, when considering children from middle childhood until early adolescence, also included increasing and decreasing trajectories. Although a stable high trajectory is often postulated (Jester, et al., 2005; Nagin & Tremblay, 1999), the data did not support the presence of this trajectory in the general population. This issue is further

discussed in chapter 4. Both the shapes of the trajectories (i.e., the initial levels of problems and the developmental course) and the proportion of children in each trajectory were similar for twins and singletons. Teacher ratings further supported the comparability of twins and singletons with respect to ADHD symptoms. This knowledge is important, especially in the light of recent findings regarding a causal relationship between ADHD symptoms and birth weight (Groen-Blokhuis, et al., 2011). In an earlier cross-sectional twin-singleton comparison, in which adolescent twins were compared to their non-twin siblings, also no differences with respect to the prevalence of ADHD were found (Ehringer, Rhee, Young, Corley, & Hewitt, 2006). However, an earlier cross-sectional Australian study found more ADHD symptoms in twins than in singletons aged 4 to 12 years (Levy, et al., 1996), and a study of 2- and 3-year-old Dutch twins found that twins showed slightly lower levels of ADHD symptoms than singletons (Van den Oord, et al., 1996). For the first time, we now confirmed the comparability of twins and singletons with respect to ADHD symptoms by a person-centered approach (i.e., growth mixture modeling). A person-centered approach focuses on differences between individuals instead of variables, and allows for differences across unobserved subpopulations. By providing a description of the development of ADHD symptoms for clusters of twins and singletons, this study gives further support to the generalizability of twin studies on ADHD(-symptoms).

Can results from twin studies be generalized to singletons?

Twins and singletons were comparable in their development of externalizing problems and of ADHD symptoms from ages 6 to 12 years. They were also comparable in their development of internalizing problems until the age of 9 years. From age 9 years onwards, twins' and singletons' trajectories of internalizing problems began to drift apart, with twins showing less problems than singletons. So, twin studies on externalizing problems and ADHD symptoms in middle and late childhood are very likely generalizable to the general population, but for twin studies on internalizing problems the generalizability is debatable in late childhood and early adolescence, but well accepted in early and middle childhood.

The generalizability of studies on internalizing problems in early adolescence in twin samples should be addressed with care, because I did not only find mean differences between twins and singletons, but also larger individual differences (i.e., larger variances) among twins than among singletons. Such results should be interpreted cautiously pending replication and further investigation of the course of these differences into adolescence and adulthood. It is intriguing that differences were found for internalizing problems only. One could hypothesize that internalizing problems are more affected by supportive sibling relationships than externalizing problems or ADHD symptoms, but this idea is not supported by earlier research (Branje, et al., 2004). Sibling influences are likely to be domain specific and further research is needed to examine the differential effects of sibling support in twin-pairs versus other sibling

pairs. The current findings support the hypothesis that twins' increased pre- and perinatal risks do not lead to higher degrees of problem behaviors (Phillips, Davies, & Robinson, 2001).

Parental divorce

Levels of internalizing and externalizing problems

Chapters 5 and 6 focused on differences between children from divorced and intact families with regard to mean levels of internalizing and externalizing problems, as well as differences in the underlying sources of individual differences (i.e, variance differences). These two studies were performed with data from the NTR and included only twins. Numerous studies over the past decades have reported that children of divorced parents show more behavioral problems than children of married parents (Amato, 2001; Amato & Keith, 1991; Hetherington & Stanley-Hagan, 1999). However, only few of these studies included children's pre-divorce behaviors, which are important to understand the direction of the relationship between parental divorce and children's problems. A unique feature of this study was that children's internalizing and externalizing problems were assessed pre-divorce (at age 3 years) as well as post-divorce (at age 12 years) using a large sample of twins.

This study yielded the following results: according to both mothers and fathers, children from divorced families had more internalizing and externalizing problems than children from intact families when these two groups were compared at the age of 12 years. However, teacher ratings of internalizing and externalizing problems did not reveal any significant differences between these two groups of children. This interesting finding may be indicative of several underlying processes, for example that children differ in the context in which they manifest their behavioral problems. It could also be that divorced parents over-report children's problems or that they are more sensitive to them as a result of their problems, such as depressive symptoms (Fergusson, Lynskey, & Horwood, 1993). However, Van der Toorn (2009) showed that maternal depressive symptoms do not bias maternal ratings of child's internalizing problems to a serious degree. Bartels, et al. (2007) reported that about 70% of the total variance in internalizing problems from ages 3 to 12 years is agreed upon by both mothers and fathers, and 1 to 17% of the total variance is rater-specific variance reflecting real behavior instead of bias or error. These rater-specific components are probably larger for teacher data.

Furthermore, there was some support for pre-divorce parent-rated externalizing problems in 3-year-old girls. This finding suggests that part of the higher levels of parent-reported externalizing problems in 12-year-old girls may be caused by factors that existed before the divorce occurred. Such factors may include parental conflict (Jenkins, Simpson, Dunn, Rasbash & O'Connor, 2005), familial stress (Booth & Amato, 1991), and family violence (Mänty-

maa, 2011). It is intriguing that girls and not boys showed pre-divorce problems, and that this effect was only found for externalizing problems. It could be the case that increased levels of externalizing problems are more notable in girls than in boys, and may therefore evoke more familial stress.

Genetic architecture of internalizing and externalizing problems

To get a more comprehensive insight into the interplay between children's problem behavior and parental divorce, genetic analyses and within pair comparisons were done with a focus on gene-environment interaction (GxE, chapter 6). We used the classical twin design to examine GxE interaction by investigating whether the amounts of genetic and environmental influences on internalizing and externalizing problems differed between children living in intact versus divorced families. At both ages (i.e., 3 and 12 years) heritabilities of internalizing and externalizing problems were slightly lower for children from divorced families than for children from intact families. However, the absolute amounts of genetic variance were similar between children from divorced and intact families, which implies that genetic effects were not suppressed by environmental factors. Heritabilities were lower for children from divorced families due to more environmental variance (which also increases the total variance). Thus, environmental factors get more important in explaining variability in children's problem behavior in the context of parental divorce. The larger total variances in children of divorced families emphasizes the great diversity of children's responses with respect to parental divorce.

The findings differed by age and by gender. Age-sensitive mechanisms include developmental changes in gene expression, and epigenetic modifications (Lenroot & Giedd, 2011). An earlier developmental twin study examining anxiety and depression from childhood to early adulthood found evidence of new genetic risk factors in early adolescence (Kendler, Gardner, & Lichtenstein, 2008). Also, genetic innovations and age-specific shared environmental influences account for some change in internalizing and externalizing problems from ages 3 to 12 years (Bartels, et al., 2004). A recent study of Vendlinski and colleagues (2011) concluded that GxE may follow a diathesis-stress pattern for some aspects of genetic risk but a bio-ecological pattern for other aspects, even when the outcome and the environmental risk factors are the same. High levels of psychosocial adversity may reduce the impact of some genes while simultaneously allowing expression of other genes that raise the liability for internalizing or externalizing problems (Vendlinski, Lemery-Chalfant, Essex, & Goldsmith, 2011). Such processes may, at least partly, explain why the genetic variances remained relatively unchanged by the experience of parental divorce (except for boys' internalizing problems). It may be too simplistic to expect the GxE effects for a given phenotype to be in

the same direction for all exposed individuals. This is also what Vendlinski and colleagues (2011) concluded in their article.

The conclusion that twins are comparable to singletons with regard to the development of externalizing problems, but not fully comparable with regard to internalizing problems, has implications for the interpretation of the results in chapters 5 and 6. As these two studies included only twins, mean levels of internalizing problems at age 12 could have been underestimated. In addition, the presumed more intimate relationship of twins as compared to other sibling pairs may have moderated the effects of parental divorce, such that twinship may be a protective factor in the association between children's adjustment and parental divorce. However, as I found no twin-singleton differences in levels of externalizing problems, twinship is not likely to bias the association between parental divorce and children's adjustment to any serious degree. The results presented in chapters 5 and 6 are innovative and valuable in emphasizing the importance of environmental influences on the development of problem behavior before and after an important life event. Replication of the reported GxE effects is necessary, and the study could be extended with adult data of these children, in order to investigate if the experience of parental divorce in childhood also affects variability in problem behaviors in adulthood.

Limitations

As data came from two different longitudinal data sets, the results presented in this thesis should be considered in light of differences and similarities between them. Important similarities of the NTR and the ZHS include the longitudinal design with assessments from early childhood to adulthood, use of the same measurement instrument (i.e., the Child Behavior Checklist) on multiple assessments, and the use of multiple raters (i.e., mothers, fathers and, teacher). Despite these essential similarities, the two samples were not fully comparable. The main differences between the samples relate to time frame, the use of different birth cohorts, response rates, selective attrition, and socio-economic status.

As the two samples are not from the same time periods, secular changes should be taken into consideration. Secular changes can include changing family structures with rising divorce rates, mothers who work more often nowadays than 20 years ago, changing ethnic distributions, overall economic growth, new media that have changed leisure activities of children, and the constantly changing and developing education system in The Netherlands. Data were used from twins born between 1986 and 1998, whereas the children from the ZHS were born between 1971 and 1979. An earlier study did not find evidence for clear secular differences in psychopathology over a 10-year period (1983-1993) (Verhulst, Van der Ende, & Rietbergen, 1997). Tick, Van der Ende, and Verhulst (2007) found increases in Dutch children's parent-reported internalizing problems over a 20-year period (1983-2003),

but these increases were very small and not consistent across age. There were no significant differences in the degrees of externalizing and internalizing problems at ages 7, 10, and 12 between twins from different birth cohorts. Furthermore, birth cohort did not predict mean scores of CBCL Attention Problems at ages 7, 10 and 12 (Derks, 2006). For these reasons, it is unlikely that secular changes or cohort effects had any meaningful implications for the results described in this thesis.

The response rates were lower for the NTR than for the ZHS, which is probably due to the different strategies that were used to retain the participants. As higher SES families may be more likely to remain in the study, the response rate difference may reflect a higher level of selective attrition in the twin sample. To overcome the difference in SES between the two samples, I took SES into account in most of the analyses. Non-response could be related to divorce status in the twin sample. Given that the observed rate of parental divorce in the NTR (10%) is lower than in the general population (20%) (Oldehinkel, et al., 2008), divorced families may be less likely to continue their participation in the NTR. The premise that parents of twins are less likely to divorce is not supported by recent research (Collins, 2011). Problematic families may be underrepresented in the NTR in comparison to the ZHS, which implies that twins' levels of behavioral problems may have been underestimated.

Finally, the NTR recruits families from any place in The Netherlands, whereas the participants of the ZHS were recruited in the province of Zuid-Holland. Because twins from the province of Zuid-Holland had the same externalizing and internalizing trajectories as twins from the rest of The Netherlands, this difference does not limit the comparability of the samples. Further, Tick, et al. (2007) showed that there were no significant differences in mean scale scores on the CBCL between children living in Zuid-Holland and children living elsewhere in The Netherlands. As The Netherlands is a small country, regional differences were not expected.

Despite their differences, the above considerations lead to the conclusion that the NTR and the ZHS provide a solid basis for a series of meaningful twin-singleton comparisons, because of the similarities in their longitudinal designs. Other limitations should be mentioned as well. In the ZHS no data are available about siblings in the families. Therefore, a number of children labeled as singletons will in fact be one of a twin-pair. Given that approximately 2% of all children are twins, the number of misclassifications was small. Nevertheless, when a pure singleton sample was used, twin-singleton differences may have been slightly larger than the differences reported in chapters 3 and 4. Perhaps, it would be more appropriate to speak of the current study as a twin-general population comparison instead of a twin-singleton comparison.

In the light of our findings with regard to the development of internalizing problems, data on sibling relationships of twins and singletons would have made a valuable contribution to this study. Also, it would have been very interesting to refine the findings on twin-singleton differences and similarities in developmental trajectories, to see if there are any subgroups of twins that deviate from the general picture, for example by distinguishing between children

Chapter 7

with low and normal birth weight or take into account maternal age. Mothers of twins are generally slightly older than mothers of singletons and problem behaviors may decrease with increasing maternal age (Orlebeke, Knol, Boomsma & Verhulst, 1998). Therefore, a statistical control for maternal age as a confounding variable in both samples would have been of interest. Unfortunately, data on birth weight, gestational age, and maternal age were not available in the ZHS. Also, a recent meta-analysis of perinatal risks in twins, showed that IVF twins had an increased risk of preterm birth and low birth weight compared to naturally conceived twins (McDonald, 2010). However, given that IVF and naturally conceived twins have the same degrees of internalizing problems, externalizing problems, and attention problems (Van Beijsterveldt, Bartels, & Boomsma, 2011), a possible difference between the ZHS and the NTR in the number of IVF conceived children is not likely to have biased the results.

There were also no data in the ZHS on family situation and parental divorce. Therefore, we could not include data from the ZHS in the studies on parental divorce. It would have been very interesting to compare the two samples with respect to the prevalence of parental divorce and children's responses to divorce. In this way, a more complete twin-singleton comparison could have been performed.

Recommendations and considerations for future research

Recommendations for future research include ideas for further studies refining and extending the current findings for twin-singleton differences, for mixture modeling and for studies of GxE interactions.

Twin-singleton differences

The generalizability of twins to singletons should be more elaborately tested. What we do not know yet is how twin-singleton differences and similarities in development of internalizing, externalizing and ADHD symptoms further develop over adolescence and into adulthood. In adolescence, self-report data may or may not confirm the findings from teachers and parents reported so far. It is worth to further study the etiology of the differences in internalizing problems that appear in early adolescence, and to investigate whether these differences were unique for a subgroup of twins. With such insights, twin studies can take the apparent limited generalizability of internalizing problems into account, for instance by controlling for the presumed moderating role of the sibling relationship. Future research should also focus on long-term trajectories (into adulthood) of twins and singletons of several different childhood psychopathologies, such as anxiety disorders.

The extended twin design, a design that includes not only twins, but also their additional siblings, can be used to further study the role of sibling relationships in twin families. Such data have been collected by the NTR from teachers of twins and their siblings and from adolescents at ages 14, 16, and 18. This enables the assessment of the magnitude of twin-specific shared environmental influences, or environmental influences specific to twin pairs that lead them to be more similar than non-twin siblings (e.g., environmental influences that siblings of the same age are more likely to share, such as having more friends in common) (Cosgrove, et al., 2011). With the above design one can test for a special twin environment for internalizing problems in early adolescence, for instance by comparing twin-twin covariances to twin-sibling covariances of internalizing problems.

Twin-singleton comparisons in which twins are compared to their non-twin siblings should, however, be distinguished from twin-singleton comparisons in which twins are compared to singletons from a general population sample. The representativeness of singletons from extended twin families with respect to child problem behaviors deserves further scientific attention, as it is uncertain to what degree extended families with twins and additional siblings are representative for families with multiple non-twin children from the general population. It would be interesting to perform a comparison of a range of child behavioral problems between 'regular singletons' (with or without additional siblings) and singletons that have twin siblings.

Mixture modeling

There is a growing interest among researchers in the use of (growth) mixture modeling techniques. In the meantime, quantitative psychologists and statisticians are still working on the validation of GMM, for instance with respect to the application of model fit criteria. When regarding studies on children's behavioral development that apply mixture modeling, it is important to consider the different ways in which researchers apply mixture modeling to their data. There is the 'simple' latent class growth analysis (LCGA), where no variation across individuals is allowed within classes, and there is the more complex growth mixture model (GMM) where within-class variation of individuals is allowed for the latent trajectory classes. In addition to these two main lines of mixture models, there can be various in-between forms. For instance, researchers may solve model convergence problems by making parameter constraints on the original GMM. As a result of this variety in mixture model applications, it could very well be the case that researchers with other opinions or preferences with regard to mixture modeling find a different mixture solution for a given data set. For instance, if I would have fit a mixture model to the Attention Problems data with variance constraints on the growth factors (i.e, no within class variation with respect to the intercepts and slopes) I would have distinguished three parallel trajectories. Without constraints on the growth fac-

tors I found three non-parallel trajectories (i.e., low, low-increasing, high-decreasing). Hence, parameter constraints can have serious consequences for the possible model solutions and subsequent interpretations and conclusions. The choice for LCGA, GMM, or an in-between form should always be based on the Bayesian Information Criterion and on model parsimony. Researchers should monitor the modeling process carefully and keep a detailed logbook of the outcomes of the different mixture models that need to be tested in order to identify the best fitting model. The modeling process includes testing and comparing models with and without covariates, as the results can be different depending on the inclusion of covariates.

As a result of the apparent lack of uniformity in the application of GMM, it may very well be the case that different researchers would come to different conclusions while working with the same data. Obviously, this is a very undesirable situation, and it emphasizes the importance of replication of any mixture modeling results. Researchers who apply GMM should be aware of their own role in the decision making process that leads to the choice of the best mixture model, and of the sensitivity and complexity of this statistical framework.

Despite these methodological issues, GMM is a valuable tool for researchers interested in individual differences in abnormal and normal development, and a promising tool for the investigation of interactions of biological and environmental processes contributing to risk and protective factors for psychopathology.

GxE interactions

There has been an explosion of interest in studying GxE interactions, and these can be studied in various ways. Because in this study genetic influences were modeled latently I found evidence for a statistical moderation by which environmental influences became more important. For researchers interested in identifying specific genes associated with external-izing or internalizing problems, our results suggest that greater efficiency in detecting asso-ciations with specific genes could be obtained by limiting samples to children who have not experienced parental divorce, because then, the relative influence of genetics is the largest.

Given that the impact of parental divorce may stretch over many years, including several years before the divorce actually occurs, future research should focus on the longitudinal impact of genetic and environmental risk on children's behavioral development. We still know too little about the impact of specific genetic risk factors over time, particularly during childhood and adolescence. Discovering such risk factors is, however, a big challenge, espe-cially given the complexity of the pathways stretching from a specific gene or environmental influence to its effects on observable behavior. Identifying changing genetic risk factors would have implications both for research (e.g., molecular genetic studies) and for treatment and prevention (Kendler, Gardner, & Lichtenstein, 2008). Also, further understanding of epi-

genetic mechanisms by which the environment affects gene expression is very important for the development of effective interventions (Lenroot & Giedd, 2011).