Young Netherlands Twin Register (Y-NTR): A Longitudinal Multiple Informant Study of Problem Behavior

Meike Bartels,[1] C. E. M. (Toos) van Beijsterveldt,[1] Eske M. Derks,[1] Therese M. Stroet,[1] Tinca J. C. Polderman,[1] James J. Hudziak,[2] and Dorret I. Boomsma[1]

[1] *Department of Biological Psychology, Vrije Univerisiteit, Amsterdam, the Netherlands*

[2] *Departments of Psychiatry and Medicine (Division of Human Genetics), Center for Children, Youth and Families, University of Vermont, College of Medicine Burlington, Burlington, Vermont, United States of America*

The Netherlands Twin Register (NTR) was established around 1987 at the Vrije Universiteit in Amsterdam, the Netherlands. The current article summarizes the longitudinal genetic analyses of maternal and paternal ratings of twins' behavior as a function of the sex of the children for the traits of aggression (AGG), attention problems (AP), anxious/depression (ANX), internalizing behavior (INT) and externalizing behavior (EXT). We found that genetic influences are the most important factor in explaining individual differences in these traits. For most phenotypes, influences of genetic factors fluctuate throughout development, with the exception of AP, for which genetic influences remain of similar magnitude. Changes in genetic influences parallel those in shared environmental influences, while nonshared environmental influences remain relatively constant. Around 10% to 20% of the variance is accounted for by parent-specific shared environment, which includes rater bias. For all phenotypes, stability throughout childhood is accounted for by genetic and shared environmental factors, while nonshared environmental influences are mainly age/measurement specific. About 15% of the phenotypic stability is accounted for by rater-specific shared environmental influences, which include rater bias. In conclusion, between ages 3 and 12 genetic factors are the most important cause of individual differences in emotional and behavioral problems.

The Netherlands Twin Register (NTR) was established around 1987 at the Vrije Universiteit in Amsterdam, the Netherlands. Many twins and their families have been invited to participate since, in a wide variety of research projects. The twin register can be divided into two parts: (1) The Young Netherlands Twin Register (Y-NTR), which focuses on twins born after 1986 (and their family members), and (2) the Adult Netherlands Twin Register (A-NTR), which focuses on young adult and adult twins born before 1986 (including their family members and spouses). All contact information comes from the same database, but data collection follows a different strategy for the young and the adult twins (for details see Boomsma et al., 2006). In this article we discuss the longitudinal study design of the Y-NTR and give an overview of results of the longitudinal studies of behavioral and emotional problems in 3- to 12-year-old children. We summarize the longitudinal results from genetic analyses for three of the syndrome scales of the Child behavior Checklist (CBCL; Achenbach et al., 1992), namely Aggression (AGG), Attention Problems (AP), and Anxious/Depression (ANX), and for the two broad band scales of the CBCL, namely Internalizing behavior (INT) and Externalizing behavior (EXT). For these scales we present the longitudinal genetic analyses of maternal and paternal ratings of twins' behavior as a function of the sex of the children. We look at the phenotypic stability of these behavior problems across a 9-year time span and decompose the covariance (stability) between ages into genetic and nongenetic parts.

Material and Methods

Sample

The Y-NTR contains data on twins from birth onwards. Twins are categorized by birth cohort and data collection is cohort driven. Nationwide data collection of all families is by mailed surveys. Parents of twins receive questionnaires when their twins are aged 1, 2, 3, 5, 7, 10, and 12 years of age. At ages 7, 10, and 12, teacher data are also collected, after written permission is given by the parents. After 20 years of research, large datasets have been obtained. An overview of the current sample sizes (as of

Received 26 May, 2006; accepted 2 July, 2006.

Address for correspondence: Meike Bartels, Department of Biological Psychology, Room 2B-47, Vrije Universiteit, Van der Boechorststraat 1, 1081 BT Amsterdam, the Netherlands. E-mail: m.bartels@psy.vu.nl

https://doi.org/10.1375/twin.10.1.3 Published online by Cambridge University Press

Table 1

Available Data at the NTR Per Age Group (Based on Returned Survey Information From November 2005)

Parental surveys	Number of twin pairs	Cohort	Response rate	Teacher surveys	Number of individuals	Cohort	Response rate
Age 1	25,147	1986–2004	—[a]	—	—	—	
Age 2	18,094	1986–2002	83%/83%[b]	—	—	—	
CBCL/Devereux[c]				TRF			
Age 3	15,053	1986–2002	73%/82%	—	—	—	
Age 5	12,054	1986–2000	67%/81%	Age 5	474	1990–1992	90%
Age 7	9025	1986–1998	63%/80%	Age 7	6451	1992–1998	78%
Age 10	6055	1986–1994	62%/80%	Age 10	5425	1989–1995	77%
Age 12	4087	1986–1993	61%/81%	Age 12	4023	1986–1993	74%
Conner's parents				Conner's teacher			
Age 7	3625	1993–1998	—[d]	Age 7	4860	1993–1998	—[e]
Age 10	3153	1991–1995	—	Age 10	3891	1991–1995	—
Age 12	2821	1989–1993	—	Age 12	2594	1989–1993	—
				Self-report	Number of individuals	Cohort	
				YSR			
				Age 12	1190	1986–1991	—[f]
				Age 14/16/18	2635	1987–1990	—

Note: [a]Official registration at the NTR becomes active after receiving questionnaire at age 1.

[b]The first response rate is the absolute response rate at each age; the second response rate is the relative response rate representing response rate if the questionnaire at the previous age is received as well.

[c]CBCL has been sent at ages 3, 7, 10, and 12, while parents receive the Devereux at age 5.

[d]Conner's and CBCL are sent in one questionnaire booklet, so response rate is overlapping with CBCL response rates.

[e]Conner's and TRF are sent in one questionnaire booklet, so response rate is overlapping with TRF response rates.

[f]No information available yet.

November 2005) is given in Table 1. These numbers change as data collection continues. Sample sizes decrease for older children for two reasons. The primary reason is that not all of the birth year cohorts have yet reached a specific age. For example, for twins born in 2000 we will not yet have data at age 10. Secondly, as with all longitudinal databases, there are families that do not participate at each wave of data collection ('drop-out'). However, they often participate again in subsequent waves ('drop-in'). Reasons for variable/episodic participation vary from parental requests to no longer take part to families moving to new addresses without notifying the registry staff. Some of these families participate again when children are older and/or when their new address has been traced.

Data Collection Procedure

Figure 1 presents an overview of the data collection at each age. At ages 1 and 2 years the survey is preferably filled out by the mother. Thereafter, data are collected from multiple raters: at preschool ages (twins aged 3 and 5), both mother and father ratings are collected; at ages 7, 10, and 12 teacher ratings are also collected; and at age 12, self-report data in subsamples of twins are obtained. At ages 14, 16, and 18 years all twins and their siblings are invited to register and to provide self-report data, if their parents consent.

Assessment of Problem Behavior

Information on problem behavior throughout childhood is obtained by continuous assessment with the Child Behavior Checklist (CBCL; Achenbach, 1991a), Teacher Report Form (TRF; Achenbach., 1991b) and Youth Self-Report (YRF; Achenbach & Rescorla, 2001). This longitudinal data collection strategy has the advantage that multiple informant assessment can be easily combined, due to overlapping items by gender, by informant, and by age. For each age group, items can be summed to form longitudinal syndrome scales and a total problem score. At age 3, seven syndrome scales are obtained (Emotionally Reactive, Anxious/Depressed, Somatic Complaints, Withdrawn, Overactive Behavior, Aggressive Behavior, Sleep Problems). At ages 7, 10, and 12, eight syndrome scales are obtained (Anxious/Depressed, Withdrawn, Somatic Complaints, Social Problems, Thought Problems, Attention Problems, Rule-Breaking Behavior, Aggressive Behavior). At all ages, two broad band scales, Internalizing and Externalizing Problems (INT and EXT) are obtained. In this article we focus on these two broad band scales as well as the syndromes Attention Problems (AP), Aggressive Behavior (AGG), and Anxious/Depressed Behavior (ANX).

Zygosity

For 1492 same-sex twin pairs, zygosity was determined by DNA analyses (n = 1103) or blood group

https://doi.org/10.1375/twin.10.1.3 Published online by Cambridge University Press

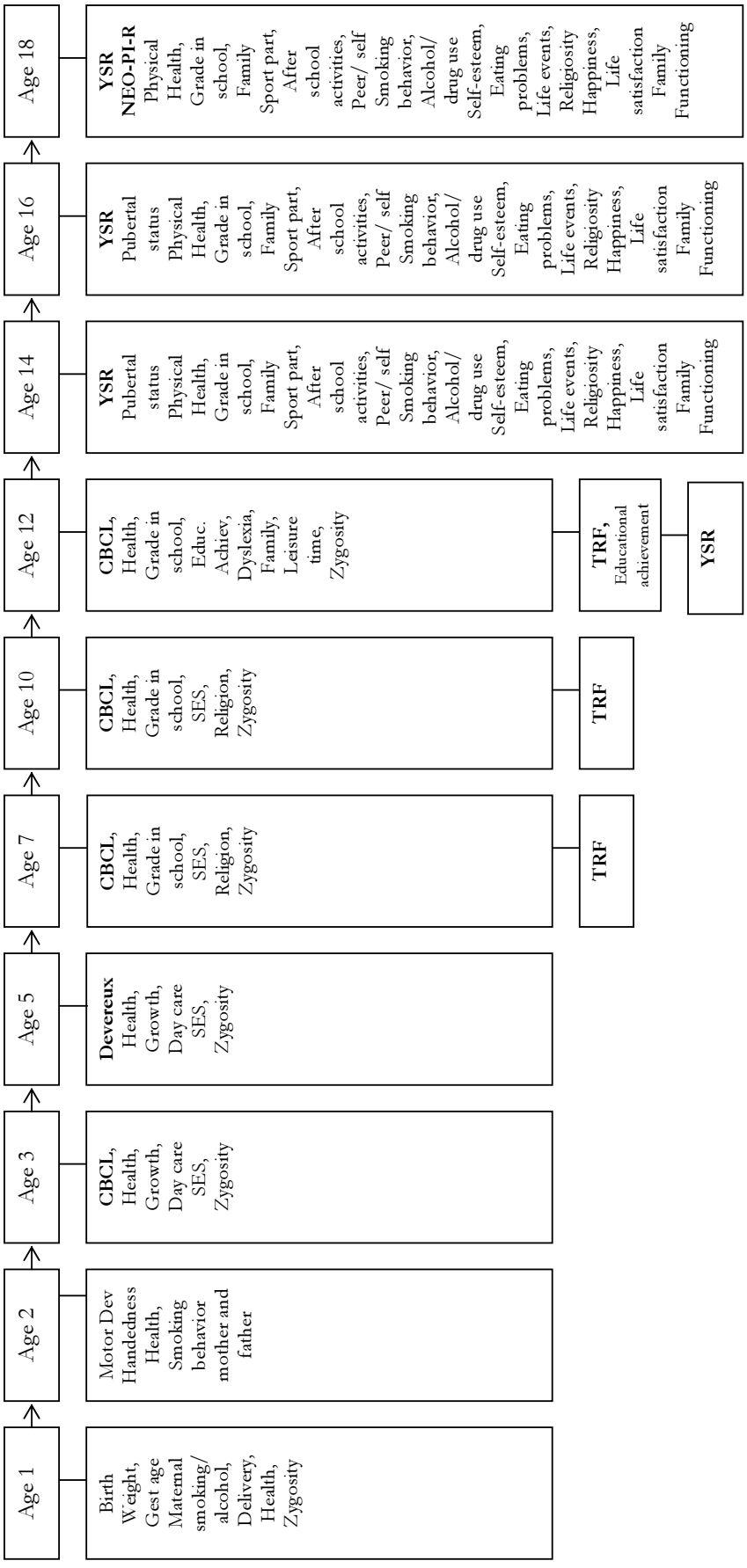

Figure 1
Overview of the survey collection of the Y-NTR.

https://doi.org/10.1375/twin.10.1.3 Published online by Cambridge University Press

Table 2

Mean Levels for INT and EXT at Age 3 for Groups With Different Degree of Longitudinal Participation

MZM		Mean INT_oldest	Mean INT_youngest	Mean EXT_oldest	Mean EXT_youngest
I.	all ages	4.39 (n = 296)	4.34 (n = 292)	17.44 (n = 296)	16.14 (n = 291)
II.	dropped out after age 3	5.22 (n = 65)	4.82 (n = 65)	20.22 (n = 65)	19.28 (n = 65)
III.	dropped out but returned to the study	5.04 (n = 46)	4.90 (n = 48)	17.20 (n = 46)	16.69 (n = 48)
	overall significance	p = .232	p = .130	p = .496	p = .071
DZM		Mean INT_oldest	Mean INT_youngest	Mean EXT_oldest	Mean EXT_youngest
I.	all ages	4.86 (n = 273)	4.32 (n = 271)	17.42 (n = 272)	15.91 (n = 271)
II.	dropped out after age 3	4.91 (n = 78)	5.15 (n = 78)	18.09 (n = 78)	17.58 (n = 77)
III.	dropped out but returned to the study	4.98 (n = 45)	5.80 (n = 46)	15.76 (n = 45)	18.15 (n = 46)
	overall significance	p = .982	p = .070	p = .445	p = .227
MZF		Mean INT_oldest	Mean INT_youngest	Mean EXT_oldest	Mean EXT_youngest
I.	all ages	4.87 (n = 360)	4.57 (n = 362)	14.97 (n = 360)	14.38 (n = 362)
II.	dropped out after age 3	5.71 (n = 66)	4.71 (n = 65)	17.76 (n = 66)	17.19 (n = 64)
III.	dropped out but returned to the study	5.37 (n = 52)	4.96 (n = 51)	18.04 (n = 52)	17.27 (n = 51)
	overall significance	p = .267	p = .806	p = .016	p = .023
DZF		Mean INT_oldest	Mean INT_youngest	Mean EXT_oldest	Mean EXT_youngest
I.	all ages	4.78 (n = 253)	4.31 (n = 254)	14.57 (n = 254)	14.34 (n = 254)
II.	dropped out after age 3	4.70 (n = 61)	4.78 (n = 58)	15.33 (n = 61)	14.56 (n = 59)
III.	dropped out but returned to the study	4.10 (n = 31)	5.10 (n = 31)	14.97 (n = 31)	13.48 (n = 31)
	overall significance	p = .654	p = .463	p = .840	p = .876
DOSMF		Mean INT_oldest	Mean INT_youngest	Mean EXT_oldest	Mean EXT_youngest
I.	all ages	4.61 (n = 269)	3.47 (n = 268)	16.41 (n = 269)	13.18 (n = 269)
II.	dropped out after age 3	4.82 (n = 98)	3.78 (n = 99)	18.02 (n = 99)	14.46 (n = 100)
III.	dropped out but returned to the study	5.00 (n = 43)	3.98 (n = 44)	16.79 (n = 43)	12.51 (n = 43)
	overall significance	p = .772	p = .382	p = .618	p = .396
DOSFM		Mean INT_oldest	Mean INT_youngest	Mean EXT_oldest	Mean EXT_youngest
I.	all ages	4.32 (n = 235)	4.15 (n = 234)	13.43 (n = 236)	14.25 (n = 233)
II.	dropped out after age 3	4.05 (n = 94)	4.55 (n = 94)	15.55 (n = 94)	16.93 (n = 94)
III.	dropped out but returned to the study	4.45 (n = 40)	4.39 (n = 38)	12.03 (n = 40)	13.87 (n = 39)
	overall significance	p = .821	p = .723	p = .066	p = .076

polymorphisms (n = 389). For all other same-sex twin pairs, zygosity is determined by questionnaire items at ages 3, 5, 7, 10, and 12. These items allow accurate determination of zygosity in 93% of same-sex twin pairs (Rietveld et al., 2000). Across the cohorts a change in zygosity distribution is observed. For cohorts 1987 to 1990 about 34% of the twin pairs are monozygotic (MZ), while for later cohorts this number drops to 27%. A possible factor may be the age of the mother: parallel to the increase of the proportion of dizygotic (DZ) twins is the significant (p = .00) increase in the average age of the mother at the time of the twins' birth. Average age is 29 in 1989, but almost 32 (31.9) in 2001. It is well known that the probability of DZ twin pair pregnancy increases with the age of the mother, due to double ovulation (Beemsterboer et al., 2006). In addition to possible associations with maternal age, the number of fertility treatments has also increased in the Netherlands, resulting in greater frequencies of DZ multiple pregnancies. No change in sex distribution of the twins is observed over the cohorts. The sample has a near 50%–50% boy–girl distribution for all cohorts, with the difference being largest (47%–53%) for the 1987 cohort.

Attrition

A matter of concern in longitudinal research is non-random drop-out. We tested whether drop-out is related to problem behavior by comparing the two broad band scales INT and EXT, at age 3 for children for three groups: *Group I* participated at all ages (3, 7, 10, and 12); *Group II* dropped out after age 3 and has not yet returned to the study; *Group III* dropped out after age 3 (at age 7 and/or 10) but returned to the study. Comparisons are made for all zygosity groups (MZ male [MZM], DZ male [DZM], MZ female [MZF], DZ female [DZF], DZ opposite-sex male female [DOSmf], DZ opposite-sex female male [DOSfm]). Data of cohorts 1986 to 1992 are used, so that data at all possible ages are

https://doi.org/10.1375/twin.10.1.3 Published online by Cambridge University Press

available. For both INT and EXT no significant differences in means between the three groups are found, except for EXT in MZF (see Table 2). Bonferonni post-hoc test revealed that, although significance is reached for this three-group comparison, the two-group comparisons (I vs. II; I vs. III; and II vs. III) did not reach significance (Oldest: I vs. II p = .086; I vs. III p = .089; II vs. III p = 1.00; Youngest: I vs. II p = .098; I vs. III p = .137; II vs. III p = 1.00). In general, it is likely that drop-out is largely random according to the definition of Little and Rubin (1987).

Analyses

Because data collection is longitudinal, the contribution of genetic and environmental influences on stability and change of problem behavior across childhood can be studied. Here, we focus on the age range from age 3 to age 12 using data obtained from maternal and paternal ratings. In order to gain insight into stability, longitudinal phenotypic correlations are calculated. Secondly, to take the multiple rater longitudinal characteristics of the data into account, a 'psychometric model' (Bartels et al., 2004; Bartels, Hudziak, Boomsma, et al., 2003; Hewitt et al., 1992) is employed and expanded for use with longitudinal data. The psychometric model allows the parental ratings to be influenced by aspects of the child's behavior perceived commonly by both parents (the common or rater-independent phenotype) and by aspects of the child's behavior that are perceived uniquely by each parent (the unique or rater-specific phenotype). In this model, both the variation of the rater-independent and rater-specific aspects can be influenced by genetic, shared environmental and nonshared environmental factors. The common latent phenotype represents that aspect of the child's behavior similarly assessed by both parents and can be considered as independent of rater biases and unreliability of the ratings. Unique perceptions of the child's behavior can arise if the child behaves differentially towards each parent or if the parents observe the child in different situations. The shared environmental effects on the unique phenotype may be confounded by rater bias, such as using normative standards or response styles. Because rater bias is independent of zygosity it may be estimated as part of the shared environmental effects. A further advantage of the longitudinal psychometric model is that nonshared environmental effects that are unique to a single measurement occasion can be distinguished from random measurement error. Random errors of measurement are age specific and are unlikely to contribute to the correlations across time in nonshared environmental effects. To expand the psychometric model to a longitudinal design we use a Cholesky or triangular decomposition (Bartels, Hudziak, van den Oord, et al., 2003; van Grootheest et al., in press). The model allows for the estimation of cross-sectional heritabilities, as well as the course of genetic influences throughout childhood, and for testing sex differences in parameter estimates. The models were fitted to data using Mx (Neale et al., 2003), by raw-data maximum likelihood estimation. This allowed the use of all twin data, whether or not there were missing parental ratings at certain time points. More technical details of genetic model-fitting analyses are reviewed elsewhere (Neale & Cardon, 1992) and additional technical information on the longitudinal multiple rater model is provided in van Grootheest et al. (in press).

Results: Summary of Our Findings on Developmental Problem Behavior

Age-Specific Variance Decomposition

For all behavioral phenotypes described in this article, genetic influences are the most important factor in explaining individual differences. For most phenotypes, influences of genetic factors (A + D) fluctuate throughout development; however, the genetic influences remain constant for AP. Changes in genetic influences parallel those in shared environmental influences (C). Nonshared environmental influences (E) remain relatively constant.

Figure 2 gives an overview of heritability estimates for INT and EXT based on the multiple rater model. Estimates are based on summation of the common and rater-independent estimates, so that, for example, the heritability estimates represent the additive genetic effects on the common view of the parents plus rater-specific additive genetic effects. For INT there are no sex differences in the magnitude of the variance components and a decrease in heritability with an increase in shared environment is observed across age. Differences in estimates are seen for mother and father ratings especially at age 3. For EXT, sex differences in the magnitude of genetic and environmental effects are found at ages 10 and 12. Further, an increase in the influence of additive genetic influences is observed between ages 3 and 7, for both raters. After age 7 the influences of genetic factors fluctuate for both boys and girls.

Figure 3 provides an overview of the variance decomposition of AGG, ANX and AP. The influences of genetic and shared environmental factors are fluctuating for AGG in both boys and girls. For ANX in boys and in girls, a decrease in genetic influence is found which is in line with the results obtained in the single rater analyses (Boomsma et al., 2005), while an increase in shared environmental influences is observed. Nonshared environmental influences on ANX remain relatively constant throughout development. AP is found to have constant and rather high influences of genetic factors. Genetic dominance is found at all ages for AP. Note that the figure provides estimates of broad-sense heritabilities (additive genetic + dominant genetic variance /total variance); however, AP is the only phenotype for which influences of genetic dominance are found. The remaining variance is attributable to nonshared environmental influences.

https://doi.org/10.1375/twin.10.1.3 Published online by Cambridge University Press

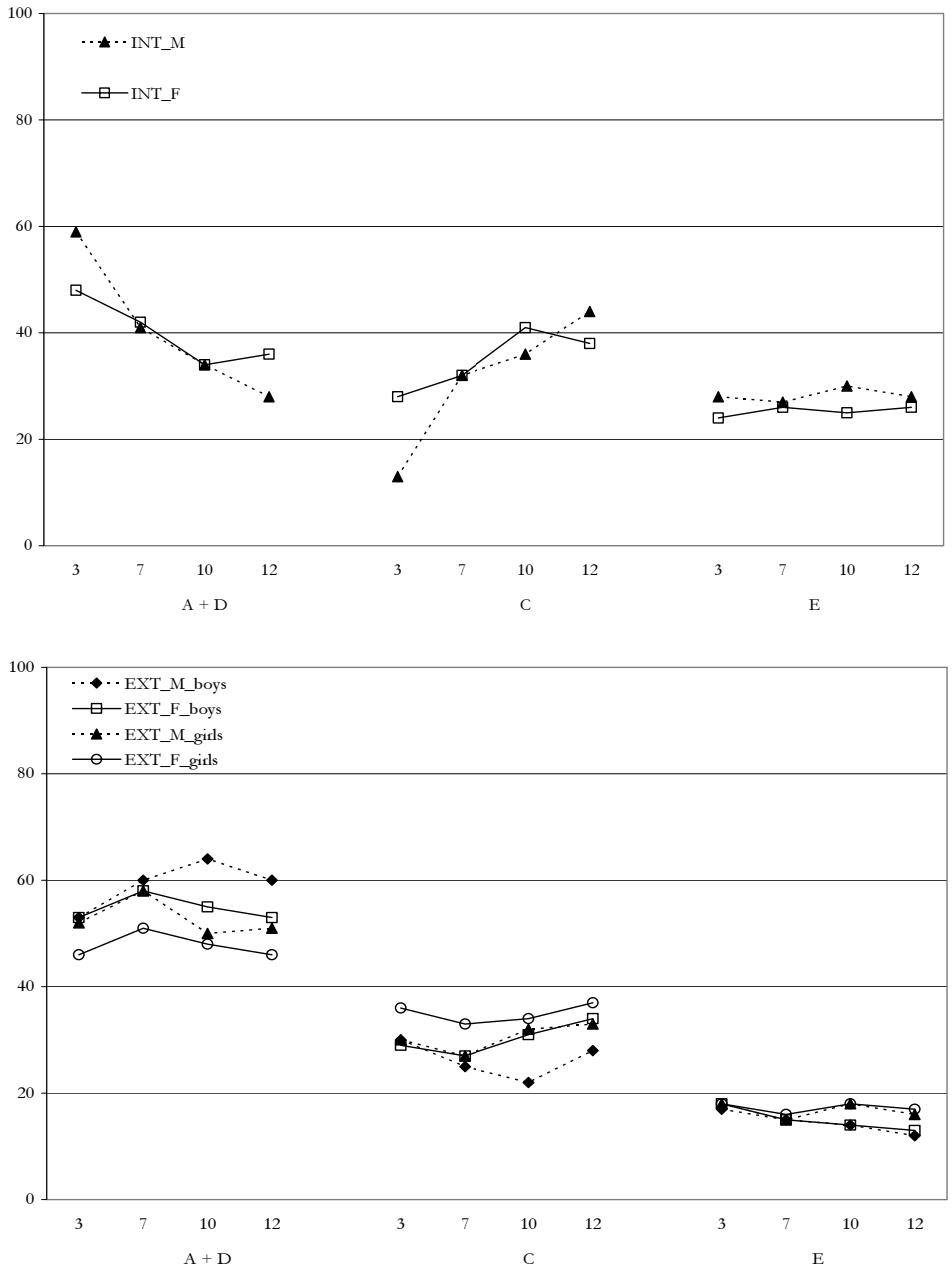

Figure 2

Variance decomposition for INT and EXT, in which M refers to mother ratings and F to father ratings.

Rater Bias and Rater-Specific Views

Because all analyses were based on mother and father ratings, they provided the opportunity to study parental agreement and disagreement. Parental correlations are in the range of .5 to .6. The phenotype rated in common shows moderate stability throughout childhood and is influenced by genetic, shared environmental and nonshared environmental influences. Influences on this common phenotype are not expected to contain rater bias or other sources of unreliability. With the multiple rater design, insight is also provided into the sources of parental disagreement. Around 10% to 20% of the variance is accounted for by parent-specific shared environment, which contains rater bias.

Stability Throughout Childhood

The longitudinal design provides insight into sources of stability. Estimates of phenotypic stability based on maternal ratings are given in Table 3. For most indices of psychopathology, there is evidence of stability, with correlations ranging from .23 for stability of ANX between age 3 and 10 to .77 for stability of AGG between age 10 and 12. Higher correlations are found

https://doi.org/10.1375/twin.10.1.3 Published online by Cambridge University Press

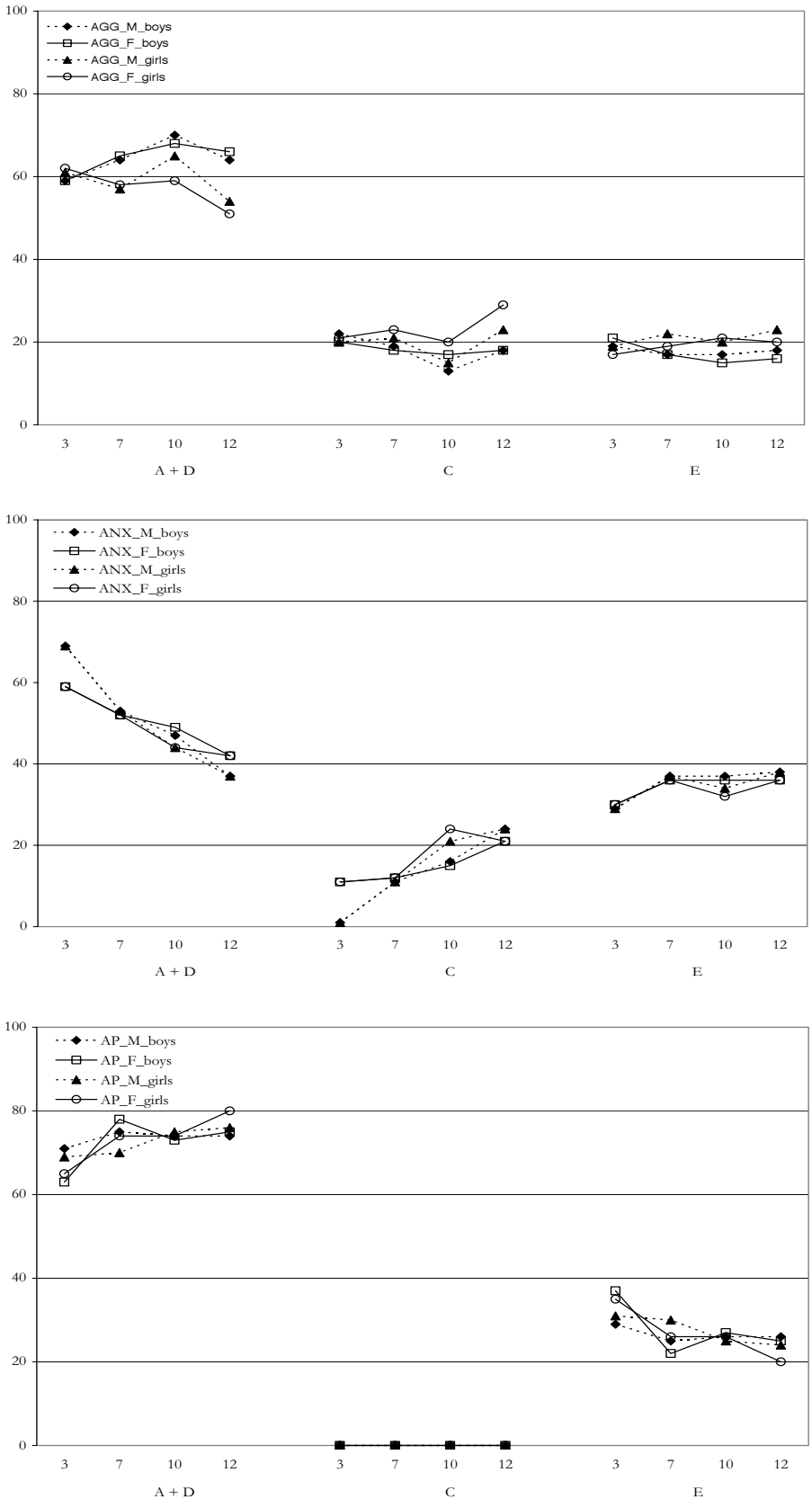

Figure 3

Variance decomposition for AGG, ANX and AP, in which M refers to mother ratings and F to father ratings.

https://doi.org/10.1375/twin.10.1.3 Published online by Cambridge University Press

Table 3

Phenotypic Correlations, Representing Stability, for Distinct Phenotypes Based on Maternal Ratings

	Internalizing			
	3	**7**	**10**	**12**
3	1	.35	.32	.29
7	.41	1	.61	.54
10	.35	.62	1	.66
12	.33	.59	.69	1
	Externalizing			
3	1	.54	.50	.49
7	.57	1	.73	.69
10	.48	.70	1	.76
12	.46	.66	.72	1
	Attention Problems			
3	1	.41	.38	.35
7	.41	1	.69	.67
10	.37	.68	1	.75
12	.38	.65	.72	1
	Aggression			
3	1	.48	.43	.42
7	.48	1	.73	.67
10	.43	.69	1	.77
12	.41	.66	.71	1
	Anxious/Depression			
3	1	.28	.23	.25
7	.29	1	.55	.47
10	.27	.55	1	.63
12	.27	.50	.61	1

Note: Boys above diagonal, girls below diagonal.

for EXT behavior and its component syndrome AGG in comparison to INT and its component syndrome ANX. For all phenotypes stability is accounted for by genetic and shared environmental factors, while nonshared environmental influences are mainly age/measurement specific. Furthermore, about 15% of the stability is accounted for by rater-specific shared environment, including rater bias.

Discussion

The longitudinal data collected by the Y-NTR have yielded some important results. For all behavioral phenotypes we described, additive genetic influences (in the case of AP, additive and dominant genetic) are the most important factors in explaining individual differences. For most phenotypes estimates of the influences of genetic factors vary across development; however, the heritability estimate for AP remains in the range of 75%. Changes in the size of genetic influences are inversely related to changes in the importance of shared environmental influences. Nonshared environmental influences remain more or less constant.

Knowledge about the etiology of the stability of behavioral problems is useful for prevention and intervention. We find that genetic and shared environmental factors are of constant magnitude. This may imply that children with a high genetic liability or children who continue to experience adverse shared environments are at greater risk for later maladjustment. For these children, a 'wait and see' policy might be considered less appropriate and an active intervention would be required.

For most of the psychopathology phenotypes we discuss, small sex differences in the magnitude of the parameter estimates were found (except for INT). Genetic influences were stronger in boys than in girls (see Figure 3). However, no evidence for sex limitation (different genes operating in boys and girls) was found, a finding that is also important for intervention and prevention. Sex-specific intervention and prevention strategies do not seem to be indicated, unless, of course, the nonshared environmental component is sex specific.

Finally, 9% to 20% of the stability in problem behavior is accounted for by unique shared environmental influences, which may in part represent rater bias. The multiple rater analyses indicate that both mothers and father are able to rate the behavior of their child and that part of the variance and covariance is rater specific. This finding is important in several ways. First, one should be aware of effects caused by the rater rather than the child, for example, paternal ratings of children both in clinical settings and in scientific research. Secondly, the high heritabilities and the effects of raters on the outcome indicate that clinical intervention for the child only may not solve all problems. Family-based intervention and prevention, in which parents are also the focus of treatment, would probably lead to more successful outcomes (Weissman et al., 2006).

In future, the longitudinal analyses will be expanded by adding teacher ratings. Teacher data have been collected over the past years and longitudinal data are accumulating. Adding data from an informant observing the child in a different setting will both provide more insight into the behavior of the child as well as the magnitude of rater-specific effects (such as bias). Furthermore, the longitudinal database will be expanded by collecting self-report data at ages 14, 16, and 18. In addition to the twins, their nontwin siblings are also invited into the study at this point. These self-report data on psychopathology, risk and protective factors, wellness and personality can be combined with the prospective data on childhood psychopathology, and other variables. The longitudinal data will be used to determine pathways into and out of adolescent wellness, psychopathology and personality. By adding the 14-, 16- and 18-year data collection waves, seven time points will become available, thus we will be able to fit genetic autoregressive models that require multiple data points.

https://doi.org/10.1375/twin.10.1.3 Published online by Cambridge University Press

Secondly, DNA collection in twins and their families is underway. Currently DNA has been collected from 1652 twin pairs (1492 same-sex twin pairs and 160 twin pairs of opposite sex) of the Y-NTR. For 704 of these twin pairs DNA is also available from their parent and, in most cases, for one or more of their nontwin siblings. Genome-wide linkage scans and genome-wide association studies are planned to localize and identify genes of interest for childhood and adolescent psychopathology.

Acknowledgments

This research has been supported by the Netherlands Organization for Scientific Research (Grant 904-57-094, Grant 575-25-006, Grant 575-25-012) and by NWO Spinoza grants to D. Boomsma (SPI-56-464), by the National Institute of Mental Health (NIMH, RO1, MH58799-03), and by the Centre for Neurogenomics and Cognition Research (CNCR) of the Vrije Universiteit, Amsterdam. Meike Bartels is financially supported by NWO (Veni; grant number 451-04-034).

References

Achenbach, T. M. (1991a). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington,VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1991b). *Manual for the Teacher's Report Form*. 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.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for ASEBA School-Age Forms and Profiles*. Burlington, VT: University of Vermont Research Center for Children, Youth, & Families.

Bartels, M., Boomsma, D. I., Rietveld, M. J. H., van Beijsterveldt, C. E. M., Hudziak, J. J., & van den Oord, E. J. C. G. (2004). Disentangling genetic, environmental, and rater effects on internalizing and externalizing problem behavior in 10-year-old twins. *Twin Research*, *7*, 162–175.

Bartels, M., Hudziak, J. J., Boomsma, D. I., Rietveld, M. J. H., van Beijsterveldt, C. E. M., & van den Oord, E. J. C. G. (2003). 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*, *42*, 1351–1359.

Bartels, M., Hudziak, J. J., van den Oord, E. J., van Beijsterveldt, C. E., Rietveld, M. J., & Boomsma, D. I. (2003). Co-occurrence of aggressive behavior and rule-breaking behavior at age 12: Multi-rater analyses. *Behavior Genetics, 33*, 607–621.

Beemsterboer, S. N., Homburg, R., Gorter, N. A., Schats, R., Hompes, P. G., & Lambalk, C. B. (2006). The paradox of declining fertility but increasing twinning rates with advancing maternal age. *Human Reproduction*, *21,* 1531–1532.

Boomsma, D. I., van Beijsterveldt, C. E. M., & Hudziak, J. J. (2005). Genetic and environmental influences on anxious/depressed behavior during childhood: A large twin study. *Genes, Brain and Behavior, 4,* 466–481.

Boomsma, D. I., de Geus, E. J. C., Vink, J. M., Stubbe, J. H., Distel, M. A., Hottenga, J. J., Posthuma, D., van Beijsterveldt, C. E. M., Hudziak, J. J., Bartels, M., & Willemsen, G. (2006). Netherlands Twin Register: From twins to twin families. *Twin Research and Human Genetics, 9*, 849–857.

Hewitt, J. K., Silberg, J. L., Neale, M. C., Eaves, L. J., & Erickson, M. (1992). The analysis of parental ratings of children's behavior using LISREL. *Behavior Genetics*, *22*, 293–317.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York; Wiley.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2003). *Mx: Statistical modeling* (6th ed.). Richmond, VA: Department of Psychiatry, Medical College of Virginia.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht, the Netherlands: Kluwer.

Rietveld, M. J. H., van der Valk, J. C., Bongers, I. L., Stroet, T. M., Slagboom, P. E., & Boomsma, D. I. (2000). Zygosity diagnosis in young twins by parental report. *Twin Research, 3,* 134–141.

Van Grootheest, D. S., Bartels, M., Cath, D. C., Beekman, A. T., Hudziak, J. J., & Boomsma, D. I. (in press). Genetic and environmental mechanisms underlying stability in childhood obsessive-compulsive behavior. *Biological Psychiatry*.

Weissman, M. M., Pilowsky, D. J., Wickramaratne, P. J., Talati, A., Wisniewski, S. R., Fava, M., Hughes, C. W., Garber, J., Malloy, E., King, C. A., Cerda, G., Sood, A. B., Alpert, J. E., Trivedi, M. H., & Rush, A. J., & STAR*D-Child Team. (2006). Remissions in maternal depression and child psychopathology: A STAR*D-child report. *Journal of the American Medical Association, 295*, 1389–1398.

https://doi.org/10.1375/twin.10.1.3 Published online by Cambridge University Press