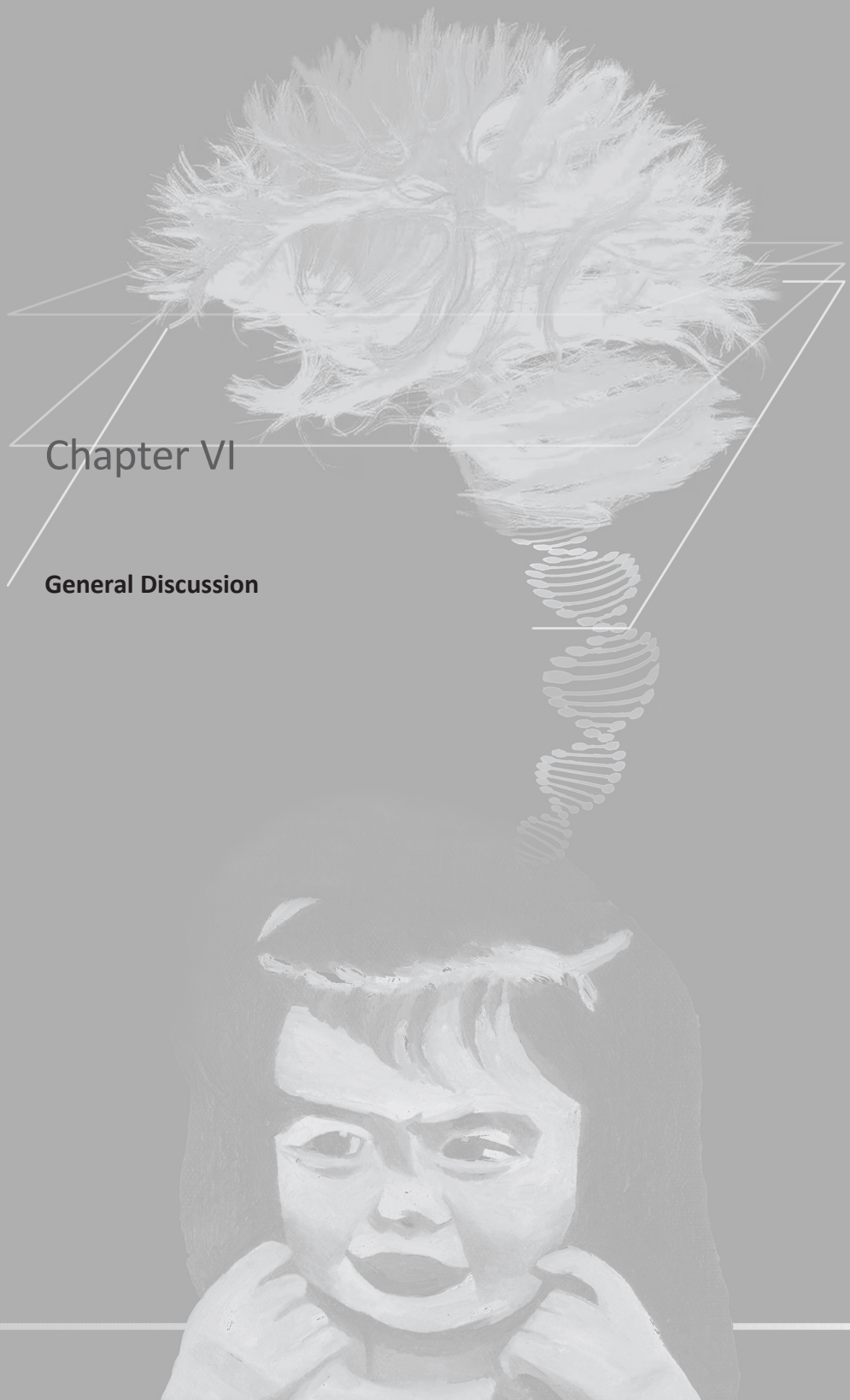

Chapter VI

General Discussion

GENERAL DISCUSSION

In this thesis I investigated the etiology and biological correlates of general psychopathology, DNA methylation in relation to ADHD symptoms and the genetics and biases of cortisol measurements. The studies confirmed some of our hypothesis, rejected others and produced some unexpected results. In the following the findings of each chapter will be summarized and interpreted. I will then discuss methodological considerations, followed by clinical implications and a thesis summary. Rather than bore the readers and repeat the discussion provided in previous chapters, I selected several findings for more in depth considerations.

Chapter III

In chapter III.A we investigated the association of parental age at delivery with child internalizing and externalizing symptoms in school-age. This study rejects the hypothesis that parental age is linked to internalizing problems. The conclusions was highly consistent across all four cohorts participating in the study. It appears that parental age either does not have adverse impacts or that the biological aspects of parental are compensated by other associated factors. A natural candidate for compensation would be better socioeconomic status (SES) with higher age, though adjustment for indicators of SES did not change conclusions. Surprisingly, higher parental age had beneficial associations with externalizing problems. Again, adjusting for SES did not change conclusions. Since the associations were only present for externalizing symptoms, one may conclude that the parental age effects are exclusive to externalizing disorders. However, without testing general psychopathology factor models, this conclusion may be premature as chapter III.5 teaches us. Using an internalizing/externalizing model without including the general psychopathology factor did not suggest any associations in that chapter, while associations with general and specific externalizing factors were detected using a general psychopathology factor model. Thus, parental age may also be associated with internalizing symptoms indirectly via general psychopathology, but future research is needed to investigate this possibility.

Chapters III.B and III.C and III.E introduced general psychopathology factors based on multiple informants in school-aged children. These models were first evaluated based on fit statistics, i.e. measures of how well they explain the observed correlation between the symptom scores. Then the psychopathology factors associated were associated with external predictors, childhood correlates and adolescence/adulthood outcomes. Specifically we tested associations with single nucleotide polymorphism, white matter integrity at age 10, school achievement test results, and with criminal behavior in adulthood, problem drinking, wellbeing and psychiatric diagnoses.

In general, we found that child psychiatric symptoms can be modeled as a consequence of both general and specific factors, as models including both general psycho-

pathology factor and specific factors (internalizing, externalizing/attention) simultaneously better explain the correlation between symptom scores compared to models of general or domain specific factors only. This finding was consistent among three cohorts of the DREAM BIG consortium. Importantly, we demonstrated that these latent factors have meaningful biological correlates and predictive validity. The general psychopathology was associated with single nucleotide polymorphisms (which explained 36% of its variance) and with less concurrent white matter integrity. General psychopathology had also predictive power, as children with higher levels of general psychopathology in childhood had a higher risk of depression in adulthood, lower well-being, more problem drinking and lower grades on high school completion exams.

At the beginning of this PhD project the bifactor models of psychopathology had been already tested in several cohorts, with the consistent result that the bifactor model has better fit than traditional models of psychopathology with correlated domain specific latent variables.[1-3] In other words, adding a general psychopathology factor to a traditional internalizing/externalizing model, better explained co-occurrence of psychiatric symptoms than only accounting for the shared variance of internalizing and externalizing factors by a correlation. However, several concerns questioned the validity of these studies. One concern was that shared method variance inflated fit statistics. Early papers, e.g.[1,2,4], fitted bifactor models to observations by a single informant. It is therefore likely that consistent ratings across all symptoms was not only due to real co-occurrence of symptoms, but also tendencies of the informant to rate higher or lower regardless of domain. Second, Bonifay & Cai[5] argue that bifactor models tend to have a bias towards better fit statistics. Fit indices do not appropriately account for the increase in functional complexity, the increased flexibility of bifactor models to fit any data independent of the number of parameters. Third, some researchers propose that bifactor models actually reflect network models.[6] In network models different symptoms can influence each other and do not necessarily originate from common factors.

In regards to the first concern about shared method variance, the risks should be reduced in this thesis, because all general psychopathology factors models were based on multiple informants and assessment time points (see methodological considerations for more further discussion). This protects against shared method variance in several ways. By having repeated measurements, in some cases with different instruments, acute situational biases affecting all symptom rating are averaged out, under the assumption that there is no systematic bias across time. An example would be a parent who is upset and frustrated about the child on the day of the assessment. This could lead to the parent rating all items indiscriminately higher without paying much attention to the specific questions asked. Ratings from multiple informants can also reduce the effect of time invariant biases, such as an overly concerned parent always rating any symptom higher independent of the behavior of the child. Higher ratings across multiple symptoms would only be taken into account if they are consistent among multiple informants. While our multi-informant approach is arguably an improvement over single-informant

designs, it also has some limitations. On the one hand, it does not protect against biases inherent to any informant, such as being more sensitive to noticing any kind of symptom if a specific symptom has previously occurred. On the other hand, the multi-informant approach may also be too stringent and adjust for unique insights each informant has, as only consistent rating are taken into account.

As for the second point, that conventional fit criteria are inadequate to account for bifactor models' functional complexity: whatever fit criteria is used, the choice of the best model cannot be judged based on statistical measures of fit alone. It is important to validate any construct with external variables, preferably with variables which can be objectively measured. In this thesis, we found that general psychopathology is associated with common SNPs, as well as with global white matter integrity. The usefulness of the general psychopathology model is especially apparent in the latter case, as white matter integrity did not associate with the traditional psychopathology factors internalizing, externalizing and attention. Thus one could have come to the erroneous conclusion that global white matter is not associated with psychopathology when analyzing psychopathology domains separately, while using a general psychopathology factor model the conclusion is that all these domains are in fact indirectly associated through general effects on general psychopathology.

The last concern, that the actual structure of psychopathology is a network and psychopathology is not the direct result of a general psychopathology factor is harder to refute. If psychopathology is best represented by network mechanisms, a bifactor model would still fit well. As an example, it is conceivable that emotional reactivity makes one more likely to react with aggression to a stressor. However, aggression may in turn lead to social repercussions, which may affect depressive symptoms, leading yet again to higher emotional reactivity. A bifactor model may suggest a general psychopathology factor underlying all three symptoms, even though in the example aggression and depression are not caused independently but result from one preceding symptom. To distinguish which model, common factor or network, better describes psychiatric disorders is highly challenging. However, even if psychopathology follows a network structure, bifactor models may still be highly useful, if a simplification. The relationship between latent variable models should perhaps be not understood as necessarily direct and independent, but perhaps should be interpreted more flexibly, as there might be mediation and feedback mechanisms between the items. While for a complete picture network modeling may be necessary, the demands for data with high temporal resolution and complexities in analysis, may make latent factor models a more practical choice for many study designs. For instance, latent variable models are well suited for genetic studies, as feedback on genetic variants from psychopathology is impossible. Furthermore, not all research question necessarily require insights whether an exposure impacts symptoms directly or via other symptoms.

In chapter III.3, we found that single nucleotide polymorphisms explained 36% of the general psychopathology factor variance (SNP heritability). To identify the specific

loci involved, we performed a genome-wide association study of a total psychiatric sum score, as a proxy of a general psychopathology factor (chapter III.4). We did not identify any specific SNPs associations in the genome-wide association study of total psychiatric problems. This does not necessarily contradict the hypothesis that SNPs are associated with general psychopathology, as the SNP hertability was 8.4%, but it does demonstrate that even sample sizes of almost 30,000 participants are not sufficient to detect SNP specific effects. Using gene-based tests, we did, however, identify the myotonic dystrophy (DM1) gene cluster as associated with total child psychiatric problems. While this locus is known for a rare mutation leading to a correspondingly rare neuromuscular disorder, common variants in this variant appear to have consequences. While common complex genetic traits and rare genetic disorders are often viewed as separates research lines, this example highlights, that both fields can potentially inform each other. Genetic correlation analyses confirmed that common psychiatric disorders co-occur in part due to shared genetic effects. Curiously, that less common psychiatric disorders such as schizophrenia or bipolar disorder co-occur due to shared genetic risk with common disorders was not confirmed. As these disorders are very rare in childhood, they were not represented well in any of the general psychopathology factor models. However, other studies in adulthood did include psychotic and manic symptoms,[2] and found that disorders, such as schizophrenia load on general psychopathology as well. It therefore appears on first glance that the genetic correlations do not follow the phenotypic co-occurrence. In other words, an individual with e.g. depression is more likely to also suffer from schizophrenia at some point in their life compared to an individual with no psychopathology. On the other hand, genetic risk for common child psychiatric symptoms appears to not predict a higher chance for psychotic symptoms. Thus, the co-occurrence of common and less common symptoms such as depression and schizophrenia must be due to a common environmental cause. However, a perhaps more likely explanation for the discrepancy is that a total psychiatric sum score in childhood is not the best measure of truly general psychopathology and needs to include measurements of thought disorder symptoms at later ages. This observation makes the high genetic correlations with other traits, such as smoking behavior, body fat and intelligence even more remarkable, as they were not included in the computation of the total psychiatric sum score. This implies, that a genetic risk for total child psychiatric problems is also a risk for many other medical problems, as the genetic risk for child psychopathology appears to also affect health related behaviors, such as smoking or overeating, which then could result in poorer mental and physical health.

Chapter IV

Next to new insights into general psychopathology, we investigated the epigenetics of ADHD examining methylation profiles at birth and school age using an epigenome-wide association study. We identified 9 genome-wide significant probes, which were differentially methylated at birth but not at school-age. Two of the probes lie in the genes

ERC2 and *CREB5*. Both are expressed in the brain and related to neural functioning or development. *ERC2* regulates neurotransmitter release and *CREB5* has important neurite growth functions. In addition, *CREB5* has been previously associated with ADHD. Due to the role in neural functioning, both probes are interesting candidates for etiologically relevant epigenetic regulations of ADHD symptoms. The association of these probes did not persist into school-age, in fact, no probe reached genome-wide significance at school-age. As discussed in Chapter IV, the overall signal of school-age methylation was lower than birth methylation, despite similar sample size. This observation was reported before for ADHD,[7] thus it seems to be a robust finding for this disorder. However, it is much less clear whether cord blood methylation has also a stronger signal for other symptoms. ADHD is a neurodevelopmental disorder with onset in early childhood. It may therefore be, that ADHD symptoms are more sensitive to early prenatal exposures, which are reflected by cord blood methylation or directly caused by perinatal DNA methylation profiles, than other symptoms. However, while other disorders, such as depression or thought disorders commonly have their incidence in adolescence or later, this does not exclude them from being influenced by early exposures. It could be that methylation profiles at birth are more important than at other timepoints for psychopathology in general. If this were the case, what are the consequences for the design of epigenetic studies and the choice of assessment periods?

On the one hand, cord blood methylation appears to have a better price to cost ratio, as under the above assumptions we can expect to find more differentially methylated regions given the same number of arrays. On the other hand, one of the promises of DNA methylation research is that changes in methylation could illuminate the biological pathways linking postnatal environment to psychopathology. This latter issue can only be addressed by repeated measurements throughout the lifespan. As DNA methylation arrays remain more costly than genetic arrays, it may for now be worthwhile to concentrate on birth methylations. Yet the case could be made that investigating biological mediation of environmental risks is of particular importance, and that this challenge should be taken up even if the bar is higher.

Chapter V

In chapter V.1 we investigated, whether polygenic scores of hair color could predict cortisol levels in hair to assess whether hair color may bias cortisol measurement using hair samples. Indeed, we observed that darker hair is associated with higher cortisol levels independent of genetic ancestry. In chapter V.2 we investigated the SNP heritability of acute saliva and cortisol measures. Neither morning levels nor total day output were related to common single nucleotide polymorphisms.

It was surprising how difficult it is to study the genetics of cortisol. Psychiatry and psychology are often criticized as being a highly subjective field with questionable unverifiable constructs. Within the field the notion exists that objectively measured endophenotypes can facilitate genetic research, as endophenotypes are “closer” to ge-

netics and thus lead to more scientific successes.[8] However, in our studies and others, subjective measures of psychopathology in most cases show consistently higher SNP heritability than 6%, i.e. the SNP heritability of acute plasma cortisol levels. Plasma cortisol were objectively measured and considered biological, yet the heritability ranged from 0 to 6% depending on the sample and method.

Perhaps more stable measures of cortisol, which are not as responsive to acute situational contexts, are needed to improve the so far somewhat inconsistent study of cortisol genetics. Hair cortisol samples are an interesting candidate for long-term profiles. However, as shown in chapter V.1, cortisol measurements in hair are biased by the darkness of hair color. This makes genome-wide analyses difficult, as researchers have to exclude SNPs in linkage disequilibrium with hair color coding genes. In addition, hair samples may not represent months-long exposures as is often claimed, but may be biased by acute events.[9] Finally, it is not a given that basal levels of cortisol are relevant for child psychiatry. In unpublished analyses, we associated hair cortisol with general and specific psychopathology factors in a structural equation model, similar to the white matter integrity models. In these analyses we were not able to detect any associations between hair cortisol and child psychopathology. Perhaps cortisol reactivity to stress is more relevant to psychopathology than basal levels, with several studies finding associations between cortisol reactivity to stressors and concurrent or later psychopathology.[10,11]

Methodological considerations

Method factors in multi-informant/method models

The use of multiple informants, methods and assessment timepoints is important to account for shared method bias and to increase precision. Shared method bias occurs when predictor and outcome are measured with the same method, in the case of questionnaire data with the same questionnaire or the same informant. This poses the danger that a bias inherent to the same questionnaire/informant is present for both the predictor and outcome, creating spurious associations, which are actually just an indication of shared bias. An example would be the association of child and parent psychopathology using parental ratings for both.[12] This problem can also affect latent factor models, as they are fitted to the covariance of two variables, which may be inflated due to shared variance biases.

However, while the theoretical advantages of using multiple informants are quite clear, the actual implementation is much less so. A simple approach is to take the average of two or more informants, however, this ignores the possibility that some items or scales are affected more by bias than others, possible resulting in overadjustment for some items and underadjustment for others. Another possibility is the use of latent factors, as used in this dissertation with two variations. In chapter III.2 we introduce latent informant factors, which load on an all items rated by the same informant. In

chapter III.3 and III.5 however, we use context factors which load on items stemming from a single rating context, i.e. from the same informant, timepoint and same instrument. These method factors explain a substantial symptom variance with either approach, typically even more than the actual psychopathology factors. For example, in the white matter integrity model, maternal ratings at age 6 had loadings between 0.39 and 0.59 on the general psychopathology factor. The loadings on the rating context factor were between 0.32 and 0.69. Interestingly, in both cases the subscale summarizing "other" items had the highest loading, suggesting that ratings on this scale are at the same time representative of general psychopathology factor but also context bias. The comparable loading between psychopathology and method factors highlights the presence of a large disagreement between raters, timepoints and instruments. It should be noted that this disagreement does not necessarily reflect the unreliability of the rater, as different raters may have different unique insights or some ages and instruments may be better suited to assess psychopathology. Therefore great care must be taken in how rating biases are corrected, as different method factors can have large impacts on the model. The models in chapter III.2 and III.3 can be easily compared as they mostly differ in how the method factors are defined. Defining the method factors on the informant level caused a bigger suppression of loadings on the psychopathology factors than defining the method factors on more narrower rating contexts (ratings from one informant with one instrument at one time-point). The first approach has a more strict control for informant effects, as these are expected to persist between instruments and time points. However, because most items in the age 6-8 years models used in III.2 and III.3 are based on mother report, there was a risk that the mother factor forms a competing general psychopathology factor instead of just covering biases. It would thus strengthen this multi-informant approach, if the number of items from the different informants were more balanced. This can be difficult though, as not all informant can rate all items. Alternatively, it may be also possible to designate the mother report as reference method, as proposed in CT-C(M-1) methods[13], and leave out the method factor for maternally rated items. All other method factors would then estimate the disagreement with the mother ratings. However, while there is an argument to be made that mothers spend most time with the children in many families and thus may have the best insight, mother ratings suffer from biases as well and therefore require correction.

Interestingly, the choice of method factor approach had great impact on the covariance between the specific factors. Using informant factors, the covariance is negligible, but with context factors the correlation is substantial. It is unclear to me where this difference stems from. One possibility is that by comprising all items from one rater across multiple instruments and time points, the method factor become a competing general psychopathology factor, reflecting the unique insight of an informant. Additionally, it may be that informant based method factors reflect the informant's general psychopathology, which in case of parental rating is related to the child's general psychopathology. Both situations imply that the specific psychopathology factors potentially are

more strictly defined in this case than if the method factors are based on single rating contexts.

In the future, it would be advantageous to have separate method factors for informants, instrument and timepoint. However, such a model with effectively four levels is likely too complex and more prone to overfitting and convergence problems. However, with a large enough sample size and with an exhaustive study design which covers all combinations of rater, instrument and time-point this may be possible. Such a design would be beneficial even with simpler method adjustments. It may be also interesting to relate the method factors to biological correlates to examine to which degree they may contain substantive psychopathology information, which is not simply bias. For example, one could explore whether parental genetic risk for psychopathology explains the parents' method factors and whether the genetic risk was transmitted to the children.

Effect sizes in epidemiology

This dissertation presents many statistically significant results, but generally low effect sizes. For instance, the top CpG site associated with ADHD explained 0.25% of the variance in the Generation R Study, global white matter explains about 0.49% of the general psychopathology variance, and maternal age at most 0.66% of externalizing behavior. At first glance, one could therefore conclude that this thesis largely identified correlates of psychopathology with little relevance and these factors are not worth considering further in the quest to explain the development of child psychopathology. The notion, that the medical field should disregard results which are statistically significant but have low effect sizes makes sense in the context of randomized controlled trials testing intervention. If a medical doctor needs to decide which treatment to prescribe, it is important to choose one which provides a meaningful change to the patient. Otherwise, the costs and risk associated with any intervention may outweigh the benefits. However, in the case of epidemiological research, where the focus is on etiology, it is much less clear what constitutes the threshold for relevance. There are no risks to weigh against and no decision of one treatment against another: many different risk factors could and probably do act jointly to cause psychiatric symptoms. This phenomenon can be demonstrated and is well accepted in psychiatric genetics. The top SNP in the GWAS of total psychiatric symptoms alone only explains 0.09% of the variance. However, the overall SNP heritability is estimated at 8.4%. This is comparable to other GWASs of continuous traits, such as depression (top SNP explained 0.03% with a SNP heritability of 4.7%) or neuroticism (top SNP explained 0.04% with a SNP heritability of 9.1%).[14] Thus while, the effects of single SNPs appear small, the effects add up to meaningful proportions.

Could the same logic be applied to non-genetic data? On the one hand, genetic variables, such as variation in SNPs, do not suffer from reverse causality biases, and confounding biases are assumed to be more limited. However, there are three notable

exceptions: population stratification, gene-environment correlations and collider bias. Population stratification occurs, when environmental influences on a phenotype differ between participants of different ancestries. Those genetic variants which differ in frequency between ancestries then are a marker for environmental differences and are thus only spuriously associated with the phenotype. Several methods are applied to adjust for population stratification, including stratified analysis, principal component adjustment or linear mixed models. Population stratification can be seen as a specific example of gene-environment correlation, though many other examples can be thought of. For instance, as parents and child share genotype, a child's genotype may also be a marker for the parenting abilities of their parents, and thus a marker for an environmental influence.[15] Finally, collider bias can occur, if selections are occuring depending on both the phenotype levels and genotype. This could e.g. occur for those participants who have both the highest genetic risk for psychopathology and higher levels of psychopathology, as they are more likely to not participate as individuals only having genetic risk or actual high levels, potentially because they have the highest burden of adverse conditions.[16]

Non-genetic variables come with an even higher risk of confounding in epidemiological studies and often contain the effects of multiple factors. Also, reverse causality is often a challenge, which can be ruled out in genetic studies. These issues are controlled for in experimental studies, so these could provide some orientation with regard to what effect sizes to expect. Previous reviews of experimental psychology studies found that the average correlation is 0.21 (SD = 0.15) between experimental condition and psychological outcome, which is equivalent to an explained variance 4.41%. When interpreting these effect sizes, one should keep in mind that experimental studies test acute effects with experimental conditions chosen strong enough (and as a consequence arguably become unrealistic and non-generalizable) to ellicit meaningful changes. This is in contrast to observational studies in general populations, where the determinants have naturally occurring distributions and outcomes are often the effects of long-term exposures. In addition, in order to control for confounding variables, determinants are adjusted for potential confounders. In the process, only the independent effects remain, but some of the effects shared with the confounder may represent true effects, which are lost. Thus, the more adjusted an analysis in an observational study is, the lower the expected effect size. This is a problem that experimental studies do not face. All in all, we therefore expect lower effect sizes in observational studies than in experimental studies. Given that experimental studies in psychology typically tend to have low effect sizes, it would be probably unlikely that we would find single causes of psychiatric symptoms explaining the majority of cases. It follows that it is necessary for non-genetic studies to identify as many determinants as possible, regardless of their effect size, and investigate their joint effect. For example, in the global white matter model of general psychopathology adjusted for maternal psychopathology, the most independently associated variable was maternal interpersonal sensitivity, explaining 3.6% of the gen-

eral psychopathology factor score variance. However, the whole model explains 28.4% (10-fold cross-validated R2, with 100 repetitions), illustrating that the aggregation of relatively few variables (19) can explain a quarter of general psychopathology. Finally, intervention strategies based on epidemiological insights may be optimized to achieve larger effect sizes compared to the naturally occurring exposure. In conclusion, research should expect that properly confounding controlled variables will have small effect sizes in epidemiological studies. However, these should not be ignored, but jointly analyzed for further insights into the etiology of psychiatric disorders.

Measurement and relational invariance of the general psychopathology factor

An important question for any study is to which populations research findings are generalizable. For instance, many of the featured studies use samples with either participants of only one genetic ancestry (e.g. European ancestry) or from only one geographical location (e.g. Rotterdam). Certainly, the most convenient scenario is when findings would be applicable to any population and are generally the same no matter of group belonging. For instance, if we could demonstrate that general psychopathology factors can be modeled the same way no matter the ancestry of participants or their socioeconomic status, then one could argue with more confidence that the findings would apply to other parts of the world, with different ethnic compositions or financial wealth. However, generalizability can be wanting in at least two ways in psychiatric epidemiology. First, different groups may express the same levels of psychopathology differently (lack of measurement invariance). It is conceivable, that the concept of general psychopathology applies to both girls and boys, but that its expression may be different. For example, boys compared to girls with the same amount of general psychopathology may show more aggressiveness (higher loadings of aggressiveness on general psychopathology) or show more aggression independent of the levels of any psychopathology factor (higher intercept of aggression). If this is the case it would be invalid to associate the general psychopathology factor across genders with any determinant or outcome, as the individual's psychopathology factors would represent different symptom sets depending on gender. In this dissertation we tested the measurement invariance of general psychopathology factors across sex, ancestry and socioeconomic status. We found that across all these groups (as defined, for example, by sex) the psychopathology factors were invariant with respect to intercepts and loadings (strong invariance), but not with regard to the residuals. Since strong invariance is a sufficient condition to associate the factors with other variables across groups, we did not discuss this issue further initially, however, I find it worthwhile diving into this matter here.

Strong invariance implies, that the basic structure of psychopathology holds across groups, and that the predicted symptom scores are the same for children with the same levels of the psychopathology factors. However, for sex and ancestry, the residual variance was not equal, which means that the factors explain different amounts of variance

depending on the sex or ancestry. Specifically, particularly externalizing and attention symptom scores are better explained in girls than boys and almost all psychiatric symptom scores have a higher variance explained in children with non-European ancestry. This in turns implies that the correlates presented in this thesis differ in the proportion of explained variance depending on the specific symptoms, the sex of the child and their ancestry. With regard to sex, a possible interpretation is that in boys externalizing symptoms are more specific, e.g. boys are more likely to show aggression without other accompanying problems, even other externalizing symptoms. This interpretation may also hold in case of observer biases, i.e. if boys do not have higher symptoms but were nevertheless rated higher, because symptoms were easier to recognize than in girls. Another possibility is, that there is another latent psychopathology factor present only in boys, though it is hard to imagine what this factor would be. In the case of ancestry, the explained variance for most subscales was higher in children with non-European ancestry. How could it be that children with European ancestry have higher proportions of variance unexplained by the tested psychopathology factors? Again, it is difficult to speculate how the missing explained proportion came about. Since the majority of the psychiatric assessments are based on parental and self report, cultural differences might play a role. Specifically, individual symptoms in children of European (mostly Dutch) ancestry could be rated higher independent of other symptoms. Rather than cultural differences per se, minority status may also play a role. Perhaps children facing more challenges due to being a minority are more likely to suffer from broader psychopathology symptoms rather than single symptoms, and thus the general psychopathology factor models fit better. Socioeconomic status as defined by maternal education does not appear to explain this discrepancy. Children with mothers from higher and lower educational background did not differ in any model parameters, including the residual error. Thus, the general psychopathology factor model is strictly invariant with respect to maternal education. Finally, the higher explained variance may be the result of higher variance in ratings of children with non-European background. It may be easier for the model to explain the co-occurrence of symptoms if there is a greater diversity in scores of children with and without psychiatric problems.

In summary, while it appears that the general and specific psychopathology factors show strong invariance across many of the tested groups, the models introduced in this thesis differ in their explanatory power. Further research is therefore needed to explore why the residual variance differs and how this could be remedied.

Clinical Implications

When drafting the chapters and presenting the results, a frequent question raised by reviewers and co-authors concerned the possible clinical implications of the findings. Indeed, at first glance the latent factors general psychopathology and the specific psychopathology factors are not observable, but abstract concepts. How would a clinician assess a patient's general or specific psychopathology levels? In the case of general

psychopathology, a good approximation would be the sum of any symptoms. However, in the case of specific psychopathology, it is more challenging to define a clinical picture. A child displaying specific psychopathology has a set of symptoms from one domain independent of their general level of psychopathology. However, independence here does not mean that both cannot co-occur. It is likely that a specific symptom occurs as the result of general and specific effects. Thus from a clinical assessment perspective, the proposed bifactor models are not directly applicable. However, these models do give valuable insights into the etiology of psychopathology and thus clues to the best prevention and treatment strategies, and may aid prediction of later psychiatric disorders. For instance, in our and other studies neuroticism or negative affect reactivity was consistently and strongly related to general psychopathology. Neuroticism is a personality trait and may be difficult to change, however, if one were to successfully change this trait, the predicted effects would be immensely helpful, as they should affect a broad spectrum of psychiatric symptoms. Alternatively, environmental stressors affecting mood would have to be elliminated as much as possible to reduce the effects of high neuroticism. On a biological level, the genome-wide and epigenome-wide association studies suggest targeting DMWD, ERC2 or CREB5 expression or its gene products, however, experimental research is needed to first confirm the causal role of these genes. Next to providing leads for intervention, biological studies of general psychopathology may be useful prediction. The combination of genetic scores, based on both general and specific psychiatric GWASs, polygenic scores taking into account gene-environment interactions, as well as polyepigenetic scores based on DNA methylation should provide meaningful predictions of psychiatric risk. The predictive power certainly will improve with increasing sample sizes of epigenetic studies.

Future research and the need for higher sample sizes

In the previous chapters specific recommendations for future research have been already discussed. Thus here I only present general observations. Increasing sample size in research of child psychiatric disorders is an important, yet challenging aspect. The need for an increase in sample size is obvious in situations where there is not enough power to detect an effect in the first place. For example, we identified three genes in the genome-wide association study of a total child psychiatric sum score, but likely many more genes are important in regulating symptoms. Likely variation at more specfic DNA methylation sites were associated with ADHD, also at school-age. While at school age no methylation levels were genome-wide significant for any CpG site, the regression coefficients did show some correlation with regression coefficients at birth in independent samples. It is therefore likely that DNA methylation at school-age is also associated with ADHD, but we do not yet have the power to detect these effectys. One may argue, that any SNP or probe with meaningful effects would have already been detected with current sample sizes and that an increase would only reveal unimportant loci. However, as discussed earlier, even very small effects seem to add up to substantial

magnitudes, thus detecting more variants and having more precise estimates of their association is of great importance.

An increase in sample size may also be important to evaluate associations for which there is enough power, but where the magnitude of the association is uncertain, such as for SNP heritability estimates presented in this study. For instance, while we can rule out large contributions of genetics towards cortisol, some genetic contribution is expected, but it is difficult to determine the precise magnitude. This is especially true for the repeated cortisol measures, for which we only had a low sample size available.

Finally, a higher sample size would allow for more sophisticated, detailed and better adjusted analysis. Even for those studies, where we had enough power to answer the main research question and have relatively precise estimates, several follow up analyses were not possible. A good example is the study on white matter. While we detected an association between white matter integrity and general psychopathology, there also appeared to be an interaction with sex. The effect appeared to be stronger in boys, but the difference was not statistically significant. It would be interesting to follow up in a larger study or meta-analysis, whether the observed sex difference was just chance or a robust effect. Also a longitudinal design examining changes in general psychopathology and white matter integrity throughout childhood would be highly interesting to determine the directionality of effect. However, such analyses may be difficult given the complexity of the bifactor model and the need to partition symptom variance into general, specific and method proportions. Convergence can become easily a problem at lower sample sizes and if there are not enough items. These problems, however, can be avoided by combining several assessment waves as done here, but for longitudinal modeling, general psychopathology measures per assesment waves are necessary, which require very large sample sizes. Furthermore, it would be also interesting to jointly analyze genetics and DNA methylation. This could help separate environmental from genetic mediation effects. In addition, interaction between the genome and methylome could be accounted for. But again, such conditional and interaction analyses need larger sample sizes than what is typically available now.

Finally, as hopefully demonstrated in this thesis, it is advantageous to analyze several phenotypes simultaneously to be able to differentiate general from specific effects and to increase precision. However, the more measures one attempts to analyze, the higher the chance that at least some data is missing. While missing data techniques can to some extent remedy this problem, higher sample sizes than in single phenotype analyses are still beneficial to account for the missingness.

Summary

In this thesis we investigated general psychopathology factors in school-aged children in relation to various biological correlates. We consistently observed across different cohorts, that a substantial proportion of variation in psychiatric symptoms can be attributed to general psychopathology effects. This observation is unlikely to be at-

tributable to informant or other rating context biases. The general psychopathology factor appears to be partly heritable, with single nucleotide polymorphisms playing a central role. One locus was identified to be associated with general psychopathology, the myotonic dystrophy cluster. Overall those genes, that are expressed in the brain, particularly in the limbic regions, appear especially important in the genetics of general psychopathology. On a neural level, more white matter across the brain is associated with lower levels of general psychopathology. At the same time, more white matter across the brain is associated with more levels of specific externalizing levels. Additionally, we observed that DNA methylation at birth is associated with ADHD symptoms in school-age and that higher maternal and paternal age is associated with less externalizing problems. Finally, we observed low SNP heritability of acute cortisol levels, but also highlight that hair cortisol levels may be biased by hair pigmentation.

REFERENCES

1. Lahey BB, Van Hulle CA, Singh AL, Waldman ID, Rathouz PJ. Higher-order genetic and environmental structure of prevalent forms of child and adolescent psychopathology. Arch Gen Psychiatry. 2011;68(2):181-189. doi:10.1001/archgenpsychiatry.2010.192
2. Caspi A, Houts RM, Belsky DW, et al. The p Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders? Clin Psychol Sci. 2014;2(2):119-137. doi:10.1177/2167702613497473
3. Laceulle OM, Vollebergh W a. M, Ormel J. The Structure of Psychopathology in Adolescence: Replication of a General Psychopathology Factor in the TRAILS Study. Clin Psychol Sci. 2015:1-11. doi:10.1177/2167702614560750
4. Lahey BB, Applegate B, Hakes JK, Zald DH, Hariri AR, Rathouz PJ. Is There a General Factor of Prevalent Psychopathology During Adulthood? J Abnorm Psychol. 2012;121(4):971-977. doi:10.1037/a0028355
5. Bonifay W, Cai L. On the Complexity of Item Response Theory Models. Multivariate Behav Res. 2017;0(0):1-20. doi:10.1080/00273171.2017.1309262
6. Nuijten MMB, Deserno MMK, Cramer AOJ, Borsboom D, … AC-C, 2016 undefined. Mental disorders as complex networks: An introduction and overview of a network approach to psychopathology. Clin Neuropsychiatry. 2016;13(4/5):68-76. https://www.narcis.nl/publication/RecordID/oai:dare.uva.nl:publications%2F5eb5a18e-1af5-429a-8ef0-f315df2bccad. Accessed January 7, 2019.
7. Walton E, Pingault J, Gaunt TR, et al. Epigenetic profiling of ADHD symptoms trajectories: A prospective, methylome-wide study Esther. 44(0).
8. Flint J, Munafò MR. The endophenotype concept in psychiatric genetics. Psychol Med. 2007;37(2):163-180. doi:10.1017/S0033291706008750
9. Kalliokoski O, Jellestad FK, Murison R. Are we overestimating the utility of hair glucocorticoids? A systematic review exploring the empirical evidence supporting hair glucocorticoids as a measure of stress. bioRxiv. July 2018:375667. doi:10.1101/375667
10. Saxbe DE, Margolin G, Spies Shapiro LA, Baucom BR. Does Dampened Physiological Reactivity Protect Youth in Aggressive Family Environments? Child Dev. 2012;83(3):821-830. doi:10.1111/j.1467-8624.2012.01752.x
11. Laurent HK, Gilliam KS, Wright DB, Fisher PA. Stress Responses : Evidence of Developmental Stress Sensitization. J Abnorm Psychol. 2015;124(1):68-79.
12. Ringoot AP, Tiemeier H, Jaddoe VWV, et al. Parental depression and child well-being: young children's self-reports helped addressing biases in parent reports. J Clin Epidemiol. 2015. doi:10.1016/j.jclinepi.2015.03.009
13. Papa LA, Litson K, Lockhart G, Chassin L, Geiser C. Analyzing statistical mediation with multiple informants: A new approach with an application in clinical psychology. Front Psychol. 2015;6(NOV):1-12. doi:10.3389/fpsyg.2015.01674
14. Okbay A, Baselmans BML, De Neve JE, et al. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. Nat Genet. 2016;48(6):624-633. doi:10.1038/ng.3552
15. Knafo A, Jaffee SR. Gene-environment correlation in developmental psychopathology. Dev

Psychopathol. 2013;25(1):1-6. doi:10.1017/S0954579412000855

16. Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol. 2017;47(1):226-235.