Genetic epidemiology of mental health

Twin and family studies of personality disorders, phobias, and symptoms of anxiety and depression

Nikolai Czajkowski

Division of Mental Health, Department of Adult Mental Health
Norwegian Institute of Public Health

Submitted for the degree of PhD at the
Department of Psychology, Faculty of Social Sciences, University of Oslo
6. DISCUSSION .......................................................................................................................... 35

6.1 Methodological considerations .......................................................................................... 35

6.1.1 Reliability ....................................................................................................................... 35

6.1.2 Validity ............................................................................................................................ 36

6.1.3 Statistical power ............................................................................................................ 37

6.1.4 Selected technical issues ............................................................................................... 37

6.1.5 Clinical samples vs. population based samples ................................................................. 38

6.1.6 Assumptions of twin and family analyses .................................................................... 39

6.1.7 Limitations .................................................................................................................... 43

6.2 Interpretation and conclusion ............................................................................................ 44

Reference List .......................................................................................................................... 49

Appendix I: DSM-IV-TR diagnostic criteria for PAPD and Cluster C PDs .........................
Appendix II: DSM-IV-TR diagnostic criteria for Phobias ......................................................
Appendix III: SCL-10 items ....................................................................................................
Appendix IV: Example R script from family analysis of SCL-10 ...........................................
ACKNOWLEDGEMENTS

This thesis presents the results of my doctoral project at the Norwegian Institute of Public Health (NIPH), Division of Mental Health. Financial support was provided partly by the National Institute of Mental Health (NIMH), and partly through the Norwegian Research Council.

The data sets that I have been privileged to have access to, and on which the papers presented in this thesis are based, have been collected over many years through the work of a large number of individuals. I would like to thank the team that initiated and administered the Mental Health in Twins study, Jennifer Harris, Einar Kringlen, Per Magnus, Ted Reichborn-Kjennerud, Liv Stene-Larsen, Kristian Tambs and Svenn Torgersen, and our international collaborators, Kenneth S. Kendler and Michael C. Neale. I am also thankful to the team behind the HUNT study, in particular former Head of the HUNT, Jostein Holmen, and the scientific staff at the HUNT subprojects at FHI, Kristian Tambs, Torbjørn Moum and Espen Røysamb. I am very grateful to be allowed to pick some of the fruits of their labour. I am also completely indebted to all the participants in the two studies.

I would like to thank my supervisors, Ted Reichborn-Kjennerud, Kristian Tambs, and Espen Røysamb, for introducing me to the field of genetic epidemiology, and for all their help and considerable patience.

I am also deeply grateful to Ken S. Kendler, Michael C. Neale, Steve H. Aggen, Kristen Jacobson and Lindon Eaves, for all the methodological advice they have given me throughout my PhD period, and especially during my visits to Virginia Institute for Psychiatric and Behavior Genetics.

I would also like to thank Arne Holte for hiring me, for his tireless and overwhelming enthusiasm, and for creating and a wonderfully stimulating environment for mental health research. Also thanks to Kari Voll and Martin Mydske Nilsen for all their help.

The division of mental health at FHI has been an amazing place to work, not least because it is solely populated by enthusiastic, knowledgeable and helpful individuals. I would in particular like to thank Ragnhild Ørstavik, Ragnhild Bang Nes, Ellen Kvestad, Christopher Nielsen, Tilman Von Soest, Eivind Ystrøm, Henrik Daae Zachrisson, Cecilie Knoph Berg, Christine Roth, Leila Torgersen, Ragnfrid Nordbø and Kjersti Gulliksen. I also want to thank Bergljot Gjelsvik for the very enjoyable months we shared an office.

My deepest gratitude goes to my family. My parents, Arja and Witold, have given me decades of nurturing and kindness, and have thought me to value and pursue knowledge. Lastly, this work would not have been possible without the love, help and support of my own small nuclear family, Alina and Sonia.

Oslo, November 2009

Nikolai Czajkowski
ABSTRACT

In this thesis we use genetically informative samples of twins and nuclear families to estimate the relative influence of genetic and environmental risk factors for a varied set of mental disorders.

In the first two papers, we consider personality disorders (PDs), a class of psychopathology characterised by marked deviations from contemporary expectations of society. To date there have been no large population based twin studies of personality disorders as assessed by structured interviews. We find that most PDs are moderately heritable, but see little empirical support for a grouping of three PDs into what is referred to “cluster C”. There was some evidence of shared environmental effects in Passive Aggressive PD, but not in the cluster C disorders.

In paper three we conduct a multivariate twin study on five kinds of phobias, to gauge the extent to which the genetic and environmental risk factors are common across the different diagnoses. We find the best model to contain two distinct liability factors, both of which are highly heritable. The first loads principally on animal phobia, while the second loads most heavily on the complex phobias, agoraphobia and social phobia. We also find that the genetic influence on blood phobia is largely unique to this disorder. For the phobias we find no evidence for common environmental influences.

In the forth paper we estimate an upper limit to the heritability of symptoms of anxiety and depression in the Nord-Trøndelag health study (HUNT). We find that these symptoms, as assessed by the ten item checklist (SCL-10), are less heritable than twin studies typically report, and we discuss possible reasons for this discrepancy.
LIST OF PAPERS

Paper I

Paper II

Paper III

Paper IV
1. CONCEPTS AND METHODS IN QUANTITATIVE GENETIC EPIDEMIOLOGY

Epidemiology is the study of factors influencing the health and illness of populations. Since being established as a scientific discipline in the second half of the 19th century, epidemiology has moved through different eras where assumptions about the underlying cause of disease, and the methods employed to find them have changed (Susser, 2006). The first epidemiologists focused on the societal level causes, such as sanitary aspects of overcrowding of urban centres due to the Industrial revolution. This was followed by a radical shift in attention towards germs and infectious diseases, when the new discipline of microbiology showed that microbes were the underlying cause of many common diseases. In the second part of the 20th century, risk factors became the primary explanatory entities in epidemiology. The concept of risk factors entails a move away from a single overarching cause of disease, to multiple factors contributing to the probability of developing a disease, none of them need be individually necessary or sufficient. Many risk factors for different mental disorder were identified, and almost exclusively they pertained to aspects of the developmental environment of the individual, such as social class, environmental adversity or family dysfunction (Regier et al., 1993; Lewinsohn, Rohde, & Seeley, 1998).

In traditional epidemiological designs the influence of genes and environment are confounded, and impossible to disentangle. Therefore, an association between an environmental exposure and a measure of mental disorder which on the surface appears to be caused by this exposure, could instead be due to underlying genetic influence common to them both. Consider for example the case of age of first alcoholic drink, and later alcohol abuse or dependence. Early onset of alcohol use has been linked to a number of negative outcomes, such as poor school achievement, behavioural problems and increased in later alcohol use (Hawkins et al., 1997). The “gateway hypothesis” posits that early alcohol use is a direct risk for the development of later abuse. However, the same pattern could be accounted for by a common set of genes predisposing to the risky non-conform behaviour characteristic both of early alcohol initiation and later alcohol abuse. It is clear that this fundamentally alters the way the results should be interpreted, as preventing early alcohol drinking under this model need not have any effect on the likelihood of later abuse. Indeed, Prescott et. al. have argued that the empirical evidence favours a common genetic liability, and not the gateway hypothesis (Prescott & Kendler, 1999).
Genetic epidemiology has been defined as “the study of the role of genetic factors and their interaction with environmental factors in the occurrence of disease in human populations” (Khoury, Beaty, & Cohen, 1993). Studies into the genetic influences on mental disorders have fundamentally impacted our understanding of their etiology. Genetic risk factors have been found for every psychiatric disorder investigated (Kendler, 2005b), and genetic epidemiology has become one of the most productive and influential approaches in the study of mental disorders.

Kendler et. al. have suggested that genetic epidemiology can be divided into four general methodological approaches or paradigms, each with their own strengths and limitations, and with the aim of providing knowledge at different levels of specificity. These are; basic genetic epidemiology, advanced genetic epidemiology, gene finding and molecular genetics (Kendler, 2005b). The papers in this thesis employ methods from the first two approaches, which are collectively referred to as quantitative genetics (QG). In essence, QG is an extension of simple Mendelian inheritance to phenotypes that do not exhibit classic recessive or dominant characteristics attributable to a single gene locus (Lander & Schork, 2006), but where numerous environmental and genetic influences are thought jointly to contribute to individual differences. Such traits are often referred to as complex or multifactorial (Hartl & Jones, 2002).

Quantitative genetic models are used to estimate the relative influence of environment and genes in determining individual differences in a given trait. While the meaning of the terms genetic and environmental may seem self-evident, within twin and family models they refer to concepts that are sufficiently abstract to warrant a short introduction. They are abstract in the sense that no specific genes or environmental exposures are typically measured, but rather these terms refer to latent and hence unobserved influences, inferred through their effect on the phenotypic similarity between different relatives.

1.1 The environment

Until the 1960, psychologists focused almost exclusively on the role of the environment when searching for developmental influences on psychopathology, and more often than not, this meant the family. Theories on how mental disorders developed implicitly assumed that offspring resemble their parents because parents provide the developmental environment (Plomin, 1989). A myriad of aspects pertaining to how families are organized, parenting styles, communication etc. have been hypothesized to constitute risk factors for development of mental disorders environment (Nichols, Schwartz, & Minuchin, 1998). If
such risk factors are characteristics of a family environment that all members share, then they ought to make all members within a family more similar, regardless of their genetic relatedness. Therefore, the goal of developmental psychopathology has been to find general classes of environmental influences that have a predictable effect in making different people exposed to them more similar in some measurable way.

In twin models, any influence that contributes positively and equally to observed similarity between relatives regardless of their genetic similarity are defined as common or shared environment (C). Conversely, unique environmental influences (E) are inferred through dissimilarity between individuals, and any proportion of sample variance that cannot be attributed to genetic or shared environmental influences is attributed to unique environmental sources. Since unique environment is a residual, any random error in the measurements will be included as part of this estimate.

Because both C and E are inferred through patterns of correlation and not measured directly, the classical twin model does not inform us regarding which specific environmental factors influence the development of a given phenotype. Perhaps the most controversial impact of quantitative genetic studies on behavioural traits has been the consistent lack of findings of shared environmental effects (Turkheimer, 2000). It is therefore important to keep in mind that a lack of significant shared environment does not necessarily mean that environmental aspects that are objectively shared by siblings, such as parental divorce, are without influence. Instead, these objectively shared events may not have an equal effect on both siblings, independent of their genetic disposition and unique environmental influences.

In extended family or twin studies, additional information is available that can be used to test for more subtle environmental influences. This includes cultural transmission, where environmental effects are transmitted from parents to offspring, or sibling specific environmental effects such as competition or cooperation (Tambs, 1999).

1.2 Genes and genetic effects

Every non sex-cell of living organisms contains within its nucleus all of the genetic information in the individual. This information is encoded in the pattern of nucleotides, which are held together by backbones of sugars and phosphate, jointly forming a double helix structure known as DNA. A gene is basic unit of heredity in a living organism. At a molecular level a gene is a region of DNA containing the information necessary to construct a protein or an enzyme, and the specific position at which a gene resides is referred to as its locus. The number of genes in the human genome is currently estimated to be between 20 000 and
The total genetic constitution of an organism is referred to as the *genotype*, while the organism’s observable characteristics are referred to as the *phenotype*. All humans have the same set of genes, but individual genes exist in alternative forms, referred to as *alleles*. When we say that a trait is *heritable*, we are implying that at least one allele has a measurable effect on the trait (Plomin, DeFries, McClearn, & McGuffin, 2001). Different genes exist in a different number of versions in the population, and different traits are influenced by various numbers of genes. Disorders caused by mutation in a single allele, are referred to as “single gene disorders”, or *monogenic*. Phenotypes such as mental disorders are believed to be influenced by a number of genes, and hence referred to as *polygenic* and *complex*.

The simplest way a set of genes can influence a trait is if each allele contributes independently to the phenotype. The total genetic influence on the trait is simply the sum of the individual contributions, and this gene action is therefore referred to as *additive* (A). Any gene action which is not additive is referred to as *non-additive*, and indicates that the effect of an allele is not independent of others. Two kinds of non-additive influences are commonly included in genetic epidemiological models, dominance and epistasis. In humans, every gene is inherited in two forms, one from each parent, and interaction between the two alleles at the same locus is referred to as *dominance* (D). If instead the influence of an allele depends on the particular pattern of alleles at other loci, its gene action is referred to as *epistatic*.

Additive genetic factors are of special interest to quantitative geneticists both for theoretical and technical reasons. They indicate the extent to which a trait will breed true, meaning the degree of parent-offspring similarity to be expected (Plomin et al., 2001). Statistical power is also highest for additive genetic effects (Eaves, 1969), which means that we are in practice often limited to investigating additive genetic effects.

### 1.3 Structural equation models

There is nothing inherent in the twin or family design that dictates the specific statistical tool to use, and a number of different approaches have been proposed. For example, heritability can be estimated by traditional multivariate regression (DeFries & Fulker, 1985), or multilevel models (Guo & Wang, 2002). However, the regression approach breaks down, once we wish to incorporate multiple (possibly reciprocally interacting) dependent variables, or multiple family relationships in the models (Heath, Neale, Hewitt, Eaves, & Fulker, 1989).

Today, nearly all twin studies use a structural equation model (SEM) approach (Heath et al., 1989). Structural equation models allow us to formulate more advanced hypotheses...
regarding the expected correlation between a set of measures, and explicitly test different models against each others. SEM is a statistical approach that can be used to test causal theories between a set of variables, and which emphasizes co-variances or correlations (Bollen, 1989). These models are often expressed in a graphical formalism referred to as path analysis. In essence, under a given structural model with a set of parameter values, an expected covariance matrix can be calculated. The parameters values are estimated by minimizing the distance between the observed and expected co-variance matrix.

A number of software packages have been developed to fit structural equation models. Mx (Neale, 2003b) is popular within the quantitative genetic community, as its syntax has been designed specifically to facilitate the implementation of multivariate twin and family models. These specialized software packages hide the complexity and many of the technical details involved in fitting structural equation models, allowing quick model implementation.

Structural equation models can also be implemented in more general purpose statistical platforms such as R (R Development Core Team, 2005). This allows for greater flexibility, gives access to a wider range of powerful statistical operations, and offers more powerful plotting functionality.

1.4 Basic genetic epidemiology

The most basic issue that genetic epidemiology can help determine about any given phenotype is whether it is familial, meaning that the phenotype clusters in families beyond what would be expected by chance. Nuclear family samples, i.e. parents and offspring, are well suited to assess the degree of familiality for a phenotype. If a significant familiality is found, the causes of the observed clustering can be investigated, and in quantitative genetics this usually entails partitioning the observed variance onto proportions attributable to the environmental and genetic influences introduced in the section 1.1.

If we limit ourselves to additive genetic influence, then abstractly, a phenotype P can be considered a function of the three influences, \( P = A + C + E \), with variance equal to \( Var(P) = Var(A + C + E) \).

Elementary probability theory lets us break the total variance into the following sums;

\[
Var(P) = Var(A) + Var(C) + Var(E) + 2Cov(A, C) + 2Cov(C, E) + 2Cov(A, E) \quad eq. 1
\]
This expression is simplified by assuming that all the covariance terms equal zero, which renders the expected phenotypic variance simply the sum of the variance attributable to the three sources.

\[ Var(P) = Var(A) + Var(C) + Var(E) \]

Narrow-sense heritability or \( h^2 \) is defined as the proportion of phenotypic variation in a population that is attributable to additive genetic variation among individuals.

\[ h^2 = \frac{Var(A)}{Var(P)} \]

On the other hand, broad-sense heritability includes all genetic influences, epistasis and dominance as well as additive effects.

Twin studies as a group have perhaps been the most influential single design in determining the relative importance of genetic and environmental influences on behavioural traits (Rutter et al., 2008). Monozygotic (MZ) twinning occurs at some stage in the first two weeks, when the zygote separates and yields two genetically identical embryos. Dizygotic (DZ) or fraternal twinning results from the fertilization of two ova by different spermatozoa, and DZ twins therefore share on average 50% of their segregating genes. As both types of twins have the same age and grow up in the same family at the same time, it is reasonable to assume that their developmental environments are very similar. These two assumptions give rise to the following expressions for the expected covariance between twins:

\[ Cov_{MZ}(P) = Var(A) + Var(C) \]

\[ Cov_{DZ}(P) = \frac{1}{2} Var(A) + Var(C) \]

\[ Var_{MZ/DZ} = Var(A) + Var(C) + Var(E) \]
These three equations are given as a path analytic model in figure 1.

While the twin design has become dominant in quantitative genetics, other genetically informative samples can also be used to estimate heritability.

Adoption studies are informative as the influence of shared environment and genes are cleanly separated. In theory there should be no correlation between the genes of an adoptee and the adoptive family, so any observed phenotypic similarity can only be due to shared environment. Frequently, adopted individuals also have siblings adopted into other families, and it is reasonable to assume that barring any contact between them, similarity between these siblings is due to shared genes.

Untangling the causes of familiality can also be done with extended family samples, where relatives beyond parents and their children are included. Unfortunately, this is often considerably more complex than using twin samples, as there is a large number of possible pedigree structures for all conceivable family constellations. Furthermore, the assumptions regarding who in an extended pedigree is subject to a shared environment are harder to justify than in the twin design. In nuclear families, familiality can be assessed, but the effect of genes and environment common to the family members cannot be separated. However, if we assume that all similarity between family members is due to genes, even nuclear families can be used to estimate an upper limit to heritability. Extended family models also have certain advantages over the basic twin models when it comes to assessing the genetic influences. This is primarily because these models allow us to test the possibility that parents are correlated with respect to the phenotype under study. While such a correlation may have a number of causes, it could be due to assortative mating, the tendency of people to have more traits in common than likely if mating was random. If the traits in question are genetically influenced, assortative mating could render siblings / DZ twins more genetically similar than expected in twin models, where it would lead to an underestimation of heritability.

In addition to having their unique strengths, adoption, family, and twin models rest on somewhat different methodological assumptions, and attaining the same estimates across these different designs serves as a way of validating the results.

1.5 Advanced genetic epidemiology

By using more sophisticated statistical models, advanced genetic epidemiology can refine the coarse estimate of heritability and environmental effects attained through the basic
approach. Methods in advanced genetic epidemiology relevant to the papers presented in this thesis are introduced in this section.

The univariate twin model can be extended to include multiple phenotypes, where covariance between the phenotypes is partitioned in a similar manner as their individual variances (i.e. A,C,E). If several different measures of mental health are included, multivariate models can be used to determine whether they share genetic or environmental risk factors. Alternatively, if the same phenotype is assessed at different points in time, multivariate twin models can be used to explore the determinants of change and stability (Neale & Cardon, 1992).

A second extension of the basic twin model is to incorporate an interaction between gender and the A,C,E parameters, the so called sex-limitation models. These models are used to test whether biological and environmental factors are involved in the etiology of the trait under study to different degrees in males and females.

1.6 Finding genes and understanding the biological pathways

Methods in genetic epidemiology extend far beyond those used in this thesis. In particular, we will limit ourselves to inferred genetic influence, and not include actual measured genes. It should be noted that the ultimate aim of genetic epidemiology is to gain as specific insight as possible into which specific genes influence a given phenotype, and understand in detail the pathways through which they operate.

Gene finding is a term used to cover approaches that not only quantify genetic influences on a disorder, but that seek to identifying specific genes or regions of DNA that are associated with an increase of risk. There are two main approaches to gene identification, linkage or association studies (Vink JM & Boomsma DI, 2002). Linkage studies search for genetic markers that are associated with variation in disease or liability in family data structures, and yields a map of low genetic resolution of chromosomal regions associated with increased risk. In association studies, high resolution set of genetic markers are compared in affected and unaffected individuals, to determine whether any of these markers are statistically overrepresented in one group.

If individual variation is found to be heritable, and specific genes are found to be associated, then ultimately it is the aim of genetic epidemiology to use molecular genetic methods to uncover the details of the pathways that lead from DNA to the abnormal brain functioning that underlies the disorder. This is still a long way off for most mental disorders, as even detecting individual genes which can account for a nontrivial amount of phenotypic
variance for measures of most mental disorders has proved challenging (Sillanpaa & Auranen, 2004). The lack of robust findings of candidate genes is at least partly believed to be due to the influence of many genes, each with small effects, rather than few genes with large effect underlying most forms of psychopathology (Kendler, 2005a).

1.7 Heritability; quantifying the genetic influence

Heritability is an abstract aggregate statistic that is easy to calculate, but somewhat harder to understand. Misconceptions about how heritability should be interpreted is no doubt partly to blame for much of the resistance with which behaviour genetic research has been met within social sciences, where it has been taken for granted that individual differences in behavioural traits are due to environmental influences.

Heritability refers to a proportion of variance. This means that the more homogeneous environmental influences are within a population, the greater the relative contribution of genes become in accounting for phenotypic variance. Alternatively, if our samples are drawn from a homogeneous gene pool, the relative measured effect of environment increases.

Since a heritability estimate pertains to a specific sample assessed at a particular point in time, it cannot strictly be generalized from one group onto another, nor to the same group at a different time. Furthermore, since variance is an aggregate statistic, heritability cannot meaningfully be ascribed to individuals. Nor does “high heritability” necessarily imply “inevitable”. Although its name is somewhat suggestive of this, “heritability” is not a measure of deterministic gene action, nor is it a hard limit of the extent to which a trait can be developed. There is nothing that theoretically precludes strongly heritable traits to be modified if sufficiently potent environmental influences can be identified. A frequently cited example of this is the autosomal recessive and hence completely heritable disease PKU, which consistently led to mental retardation and IQ often below 50, before diets low in phenylalanine were found to be an effective treatment (Plomin, 1989).
2. MENTAL DISORDERS

2.1 The Diagnostic and Statistical Manual of Mental Disorders (DSM)

The DSM-IV (American Psychiatric Association, 1994) ranks alongside the largely overlapping ICD-10 (Dilling, Mombour, & Schmidt, 1991) as the most widely used nomenclature for classification of mental disorders. The first edition of DSM was published in 1952, but it was not until the third edition (1980) that it took its current modern form. Previously broad and etiologically defined categories continuous with normality received strict operationally defined criteria of symptom based categorical diseases (Mayes & Horwitz, 2005). The aim of DSM-III was to become more theory-neutral, as pathogenic processes were no longer used to organize categories. This allowed clinicians to communicate a common set of categories, even if they had different views regarding etiology. When observable symptoms were used to define a disorder, standardized instruments like structured interviews could be developed. This increased diagnostic reliability, and made psychiatric disorders easier to study empirically. The impact that operationally defined criteria has had on the epidemiology of mental disorders is illustrated by the US-UK study (Cooper, 1972). Until the early 70’s, prevalence estimates were chiefly based on hospitalization records, which seemed to suggest a substantial difference in prevalence between the US and UK across a range of mental disorders. The US-UK study was the first large epidemiological study in which psychiatrists were trained in the use of structured interviews, and where a common set of diagnostic criteria were applied. This resulted in similar estimates of prevalence of mental disorders across the countries, suggesting that previously reported differences were due to varying diagnostic practices rather than genuine differences in morbidity patterns (Tsuang & Tohen, 2003).

In the DSM, a mental disorder is defined as a clinically significant behavioural or psychological syndrome or pattern that occurs in an individual and that is associated with present distress or disability, or with a significantly increased risk of suffering death, pain, disability, or an important loss of freedom (American Psychiatric Association, 1994). While this definition emphasizes a clinical cut-off between normality and pathology, in this thesis we will also consider measures of mental health in the sub-clinical range.

The DSM-IV uses a multi-axial approach, which entails that diagnoses are made relative to five aspects of disorder or disability, the two first being the principal ones. This subdivision was first introduced in the third edition to encourage clinicians to observe and code aspects of functioning that may otherwise be overlooked in the presence of the more
salient symptoms. Axis-I encompasses the clinical disorders, such as depression, anxiety and schizophrenia, while the axis-II is designated for coding personality disorders (PDs) and mental retardation. PDs are defined as “an enduring pattern of inner experience and behaviour that deviates markedly from the expectations of the culture of the individual who exhibits it” (American Psychiatric Association, 1994). These behavioural patterns are severely inflexible and persistent, and often cause serious personal and social difficulties. Placing the PDs on a separate axis was motivated by a need to increase their clinical visibility and their importance as an object for research. The distinction was also partly based on the prevailing understanding of important differences between these two classes of disorders, as PDs were believed to be more lifelong and pervasive, to develop at a younger age, and be more resistant to treatment. However, there is no strong empirical basis for either of these assumed differences (Krueger, 2005).

The 10 PDs included in the main body of the DSM-IV are grouped into three clusters, the “odd-eccentric” cluster A, the “dramatic-emotional” cluster B, and the “anxious-fearful” cluster C. These clusters were first introduced in DSM-III and retained in later editions, even though they have been criticized for lacking in rationale and have received mixed empirical support (Schopp & Trull, 1993). This is acknowledged by the DSM, which emphasizes that the classification of PDs into these three clusters is based on ‘descriptive similarities’, has ‘serious limitations and has not been consistently validated’ (American Psychiatric Association, 1994).

2.2 Mental disorders; continuous or discrete

The fundamental organizing principle in the DSM is the syndrome, a pattern of symptoms appearing together temporally in different individuals. For each syndrome, the DSM lists a collection of symptoms, and a diagnosis is warranted if an individual meets more than a set number of criteria. The DSM therefore represents mental disorders as categorical entities. The reasons for choosing a categorical approach to the diagnoses of mental disorders were complex, based not only on the existing scientific evidence, but also by health-political considerations such as the need for clear distinctions between cases and non-cases (Mayes & Horwitz, 2005). The limitations of a categorical approach are clearly acknowledged by the DSM manual itself, and clinicians are encouraged not to equate a diagnostic category with a disease. Still, the problems associated with a categorical view have mounted to the extent that the DSM-V workgroup will consider dimensional models for inclusion in DSM-V, scheduled for publication in 2012 (Regier, 2007).
First, there is the problem of the considerable heterogeneity within diagnostic categories. A frequently cited example is obsessive-compulsive personality disorder, where two individuals can meet the diagnostic requirement without sharing a single overlapping criterion (Widiger & Trull, 2007). Therefore, individuals classified in the same diagnostic category can have very different diagnostic and prognostic profiles.

Second, there are unclear boundaries between disorders, and excessive comorbidity between disorders that are assumed to have distinct causes. After publication of DSM-III, it was hoped that research would confirm the proposed categories by identifying distinct etiologies for the different disorders, and that recommended treatment similarly would follow categorical boundaries, but this has not happened. Instead, and as will be discussed in more length later, multivariate twin studies have indicated that both genetic and environmental risk factors are often common to several disorders. Similarly, it has become clear that there is a lack of specificity in treatment response, and medications such as the SSRIs have proved effective for a wide array of different mental disorders, such as major depression, social anxiety and borderline personality disorder (American Psychiatric Association, 1994). The problem of comorbidity is perhaps the greatest challenge to address in DSM-V, and will be discussed more in depth in the next section.

Third, boundaries between pathology and normality have been criticized as scientifically arbitrary (Widiger & Trull, 2007). Critics argue that such boundary disputes may be the result of arbitrary distinctions being imposed on an underlying continuous domain of functioning, and the arbitrary boundaries exist not only between normality and pathology, but also between different disorders.

Recently, taxometric methods have been developed to more formally test whether mental disorders are best considered categorical or dimensional. Several studies have reported that certain Axis-I disorders such as major depression may best be represented as dimensional (Ruscio & Ruscio, 2000). However, the categorical framework has been argued to be especially problematic for the Axis-II disorders, and Widiger goes so far as to say that “There is little doubt that someday the classification of personality disorder will be dimensional” (Widiger, 2007). While a move to dimensional representations is likely, at least for axis-II disorders, what form this model is to take is still highly uncertain (Trull & Durrett, 2004), and by one count, there are 18 alternative proposals for a dimensional classification of personality disorder alone (Widiger, 2007).

The issue of dimensional models of psychopathology is highly relevant to this thesis, as all the papers included are based on analyses of measures of mental health that are
dimensional rather than categorical. In paper I and II, we have used a sum-score of threshold and sub-threshold axis-II symptoms. While the majority of the phobia diagnoses in paper III are full threshold diagnoses, we have also included sub-threshold scores, where all but one of the DSM-IV criteria are endorsed. Finally, in paper IV, the family analysis of symptoms of anxiety and depression, we leave the DSM altogether and investigate the heritability of self-reported symptoms of anxiety and depression.

2.3 The prevailing problem of comorbidity

The term *comorbidity* was originally coined by Feinstein (1970) with reference to chronic disease, as “any distinct additional clinical entity that has existed or that may occur during the clinical course of a patient who has the index disease under study” (Feinstein, 1970). In epidemiological literature it is common to distinguish between concurrent and lifetime comorbidity. *Concurrent comorbidity* refers to two or more disorders being present at the same time, while in *lifetime comorbidity* the disorders are not necessarily overlapping in time. Quantitative genetic studies typically limit themselves to studies of lifetime comorbidity, as it cannot be assumed that a genetic influence should manifest in a disease in two related individuals within a narrow window of time.

Difficulties became apparent when the original definition of comorbidity was applied to mental disorders, as it was necessary to specify what is meant by a *distinct* clinical entity. It was quickly recognized, and now widely acknowledged that with mental disorders, concurrent comorbidity is the norm rather than the exception, with the rate dramatically increasing if lifetime comorbidity is assessed (Widiger & Samuel, 2005). The *problem* refers to the fact that this high level of co-occurrence is hard to reconcile with the current understanding of mental disorders as distinct, each with its own etiology and pathology. The excessive levels of comorbidity between mental disorders are perhaps most convincingly demonstrated in the DSM-III-R based National Comorbidity Survey (NCS) (Kessler et al., 1994) and the DSM-IV based National Comorbidity Survey-Replication (NCS-R) (Kessler, Berglund, Demler, Jin, & Walters, 2005). These large population based studies have also convincingly shown that the comorbidity observed in clinical studies are not solely an artefact due to sampling errors (Krueger & Markon, 2006).

In the NCS-R, it was found that 45% of individuals meeting the criteria for one disorder the past 12 months also met the criteria for at least one other during the same time period. Furthermore, more than half of all lifetime diagnoses occurred in 14% of the
population with a history of three or more comorbid disorders (Kessler et al., 1994). Comorbidity is high within clusters of related disorders, such as the anxiety disorders. Curtis (1998) found that nearly 76% of individuals with a lifetime DSM-III-R simple phobia reported one or more other co-occurring phobias.

Comorbidity is particularly high among the Axis-II disorders (Oldham et al., 1992; Grant et al., 2004; Stuart et al., 1998).

Lastly, comorbidity is also high across axis-I and II, a further indication of the arbitrary separation of personality disorders from the axis-I disorders (Lenzenweger, Lane, Loranger, & Kessler, 2007).

Epidemiological research has firmly established that the observed comorbidity is not merely an artefact of overlapping classification criteria or recruitment bias, but is instead a result of many factors (Klein & Riso, 1993). Twin samples are well suited to determine the cause for comorbidity, as models that are indistinguishable in ordinary cross-sectional designs can be distinguished when considering twins. Neale et al. have developed twin models to assess twelve different models of comorbidity (Neale & Kendler, 1995). For example, the occurrence of one disorder can increase the risk for another disorder, or two disorders can represent alternate forms or manifestations of the same underlying liability. The few studies to investigate these questions by means of quantitative genetic methods offer tentative results to the effect that many risk factors for psychopathology are not disorder-specific, and that a large number of psychiatric disorders may be explained more parsimoniously by a small number of underlying factors (Kendler et al., 1995a). The tentative conclusion that can be drawn from the existing body of evidence, lends most credence to correlated genetic and environmental liabilities as an explanation of the comorbidity in psychopathology (Krueger & Markon, 2006; Middeldorp, Cath, Van Dyck, & Boomsma, 2005).

2.4 The quantitative genetics of mental disorders

The volume of literature produced on the quantitative genetics of psychopathology allows for only a brief outline of the major contributions of this field. Twin and family analyses have profoundly affected our understanding of the etiology of mental disorders in several ways;

First, genetic risk factors have been found for every psychiatric disorder investigated (Kendler, 2005b). For Axis-I the highest heritability estimates are usually found for schizophrenia, where meta-analytic studies estimate a heritability of roughly 80% (Sullivan, Kendler, & Neale, 2003). Estimates of heritability for bipolar affective disorder cluster around
75% (Smoller & Finn, 2003), for autism around 90% (Freitag, 2007), and for attention-deficit/hyperactivity disorder 76% (Biederman & Faraone, 2005). Disorders displaying moderate heritability in the range 25%-45% encompass both most kinds of anxiety disorders (Hetteja JM, Neale MC, & Kendler KS, 2001), major depression (Sullivan PF, Neale MC, & Kendler KS, 2000), and alcohol abuse and dependency (Walters, 2002). Analyses based on clinical samples suggest that most PDs also are moderately heritable (Torgersen et al., 2000). Furthermore, the close association between normal and abnormal personality functioning (Markon, Krueger, & Watson, 2005), and the support for the etiological role of genetic factors in normative personality traits found in family, twin and adoption studies (John, Robins, & Pervin, 2008), also gives ample reason to expect genetic influence on PDs.

Second, while a substantial heritability has consistently been found for all mental disorders, shared environmental effects are rarely reported. This is in stark contrast to traditional psychological theories of etiology in which only environmental effects are included, and suggests that many of the environmental influences that are associated with mental disorder have no effect on their development that is independent of the genetic constitution of the individual.

Third, genetic epidemiology has impacted mental health research by finding convincing evidence that comorbidity between many mental disorders is largely due to a common genetic influence. For example, a recent study by Lichtenstein et. al., in which a multi-generation register of more than two million nuclear families was analysed, found that 63% of the comorbidity between schizophrenia and bipolar disorder best was explained by additive genetic influences. Anxiety and affective disorders are also widely recognized to be highly comorbid (Maser & Cloninger, 1990). Quantitative epidemiological analyses of generalized anxiety disorder (GAD) and major depression have implicated a genetic correlation in the range 0.86 to 1.0 (Kendler, 2004; Roy, Neale, Pedersen, Mathe, & Kendler, 1995). Common genetic and environmental liabilities are, as discussed in the previous section, the explanation of comorbidity with the most empirical support.

Fourth, many measures of mental health show pronounced gender differences in prevalence. There is now substantial evidence indicating differences in genetic risk factors for major depression, and higher levels of heritability in females (Bierut et al., 1999; Kendler K.S., Gatz, , Gardener, & Pedersen, 2006).
2.5 Areas in need of more study

While genetic epidemiology has contributed considerably to the understanding of the etiology of mental disorders, there are areas where current knowledge is sparse. While enough twin studies have been conducted on all major axis-I disorders to get robust meta-analytic estimates of heritability (Sullivan PF et al., 2000; Hettema JM et al., 2001), estimates for axis-II disorders are almost completely lacking (Reichborn-Kjennerud, 2008). Indeed, the Axis-I/Axis-II is described in detail in section 4.1 is the first quantitative genetic analyses on axis-II disorders that has been carried out on a population-based sample using a standard interview-based instrument. The need for empirical study is especially important for the axis-II disorders in the appendix of the DSM, passive aggressive PD (PAPD) and depressive PD, for which the DSM workgroup has explicitly called for more empirical studies before decisions regarding their status in future versions can be settled (American Psychiatric Association, 1994). The first paper in this thesis therefore aims to investigate the heritability of PAPD in a population based sample.

Since even the most basic genetic epidemiology is lacking for PDs, it follows that scarcely any multivariate twin analyses have been conducted on these disorders. We therefore have little understanding regarding the relative influence of genes and environment in accounting for the comorbidity between PDs. This constitutes a considerable gap in psychiatric literature, as the comorbidity between PDs is consistently found to be among the highest of all mental disorders. In our second paper we therefore conduct a multivariate twin study on the cluster C PDs to investigate the cause of their lifetime co-occurrence.

Furthermore, multivariate twin analyses are also underutilized for disorders where numerous univariate studies have been performed, such as anxiety subtypes. In our third paper we consider one such class of disorders, phobias, and present results from the first multivariate twin study to include all DSM-IV phobias in a sample of female twins.

Even for the disorders where multivariate analyses have been published, independent replications are necessary, as individual studies are often severely underpowered (Neale, Eaves, & Kendler, 1994). It is also necessary to replicate results using a different methodological design, as each approach rests on a particular set of assumptions that may be violated. Specifically, twin studies have become so dominant in quantitative genetics that it would benefit the field if these results to a greater extent would be replicated in samples of families. This is particularly true given the frequently cited, but scarce evidence that family and adoption studies generally yield lower levels of heritability than do twin studies (Thomas J.Bouchard & John C.Loehlin, 2001; John et al., 2008).
3. RESEARCH OBJECTIVES

Paper I
To investigate the familial aggregation of passive-aggressive personality disorders (PAPD), and explore other issues regarding this disorder raised by the DSM-V Personality Disorder Work Group.

Paper II
To study the genetic epidemiology of the DSM-IV cluster C personality disorders, and examine the validity of the cluster C construct by determining to what extent common familial factors influence the individual PDs.

Paper III
To examine, using multivariate twin analyses, the structure of the genetic and environmental risk factors underlying lifetime comorbidity of DSM-IV phobias.

Paper IV
To estimate an upper limit on the heritability of symptoms of anxiety and depression in a large population-based nuclear family sample.
4. MATERIALS AND METHODS

The papers presented in this thesis are based on analyses of two separate samples, the, Norwegian Institute of Public Health (NIPH) Twin Panel, and the Nord-Trøndelag Health Study (HUNT-2).

4.1 Norwegian Institute of Public Health Twin Panel, Axis-I/Axis-II study (AI/AII)

The Norwegian Institute of Public Health in Oslo has a population-based twin panel referred to as The NIPH Twin Panel (Harris, Magnus, & Tambs, 2002a). The current panel includes information on 15,370 like- and unlike-sexed twins born from 1967-1979. The database includes information from the Norwegian Medical Birth Registry (MBR), longitudinal questionnaire data, DNA, and information collected in a number of clinical sub-studies. The twins are identified through information about multiple births contained in MBR. The MBR was established January 1st, 1967, and requires mandatory notification of all live- and stillbirths of at least 16 weeks gestation. A total of 15,370 twins were born in Norway during the 13 years from 1967 to 1979. During that time period, the proportion of pairs in which both twins survived to age 3 ranged from 82 to 89 percent. The twins from these intact pairs are recruited into the NIPH program of research through mailed questionnaires. Two questionnaire studies have been conducted thus far, Q1 in 1992 (twins born 1967 – 1974) and Q2 in 1998 (twins born 1967 – 1979). Altogether, 12,700 twins received the second questionnaire, and 8045 responded after one reminder (response rate 63%). The sample included 3334 pairs and 1377 single responders.

Data for the current study derives from an interview study of axis I and axis II Psychiatric Disorders, which was carried out between June 1999 and May 2004. Participants were recruited among the 3153 complete pairs who responded to the second questionnaire and agreed to participate in the interview study, and 68 pairs who were drawn directly from NIPHTP. Altogether 2794 twins (44% of those eligible) were interviewed for the assessment of PDs. The mean age of participants was 28.2 years (range 19-36). A summary of the data collected on the NIPH twin panel is given in figure 1.
Zygosity

Zygosity was initially determined by questionnaire items previously shown to categorize correctly 97.5% of pairs (Harris, Magnus, & Tambs, 2002b). In all but 385 like-sexed pairs, where one or both of the twins was either unwilling or unable to donate a blood sample, zygosity was also determined by molecular methods based on the genotyping of 24 microsatellite markers. Discrepancy between classification based on questionnaire and DNA markers was detected in 12 MZ pairs and 5 DZ pairs (2.51%), implying an expected misclassification rate of 0.67% for the whole sample. The sample consists of 1,022 males and 1,722 females; 221 monozygotic male (MZM) pairs, 116 dizygotic male (DZM) pairs, 448 monozygotic female (MZF) pairs, 261 dizygotic female (DZF) pairs, 340 dizygotic opposite sex (DZO) pairs and 22 single responders.

Interviewers

Interviews were conducted face-to-face except for 231 interviews (8.3%) that for practical reasons had to be done over the telephone. Interviewers were mostly psychology students in the final part of their training and experienced psychiatric nurses, trained by professionals (one psychiatrist and 2 psychologists) with extensive previous experience with the instrument. All received a standardized training program by teachers certified by the WHO and passed a user license test for the CIDI. They were followed up closely individually.
during the whole data collection period, and regular meetings were also held with all interviewers present to discuss potential problems. Each twin in a pair was interviewed by different interviewers blind to the results of the co-twin.

4.2 The Nord-Trøndelag Health Study 1995-97 (HUNT-2)

From August 1995 to June 1997, the population aged 20 years or older of the 24 municipalities of Nord-Trøndelag County, Norway, was invited to take part in a health screening survey, the Nord-Trøndelag Health Study (HUNT-2). The survey included as an integrated project the Nord-Trøndelag Hearing Loss Study (Tambs, Borchgrevink, & Samuelsen, 2003) and the populations of 17 of the 24 municipalities were invited to participate in the hearing loss study. As part of this study, the participants completed a questionnaire containing the SCL-10, a shortened version of SCL-25 (Hesbacher, Rickels, Morris, Newman, & Rosenfeld, 1980) designed to measure symptoms of anxiety and depression. Valid SCL-10 scores were registered on 46,064 individuals, 21,696 males and 24,368 females. The mean age in the sample was 48.5 years, (48.8 for males, 48.8 for females). The participation rate was 68.7 percent, 64.7 percent among males and 72.7 percent among females for all municipalities together except for one, in which for certain reasons only 42.1 percent participated.

First-degree relationships were obtained from registries administered by the governmental agency Statistics Norway, identifying mother-offspring pairs with absolute certainty but with a slight chance that the father registered at birth is not the biological father. In addition to first degree relatives, data identifying spouses were supplied.

4.3 Sample

Paper I and II

In papers I and II, data from all 5 zygosity groups from the AI/AlII sample were included in the analysis. The sample consequently consists of 1,022 males and 1,722 females; 221 monozygotic male (MzM) pairs, 116 dizygotic male (DzM) pairs, 448 monozygotic female (MZf) pairs, 261 dizygotic female (DZf) pairs, 340 dizygotic opposite sex (DZO) pairs and 22 single responders.

Paper III
Female twins who had responded to the axis-I interview were selected from the AI/AII sample, resulting in an effective sample of 710 complete twin pairs (446 female monozygotic (MZ) and 264 female dizygotic (DZ)), and 10 single responders.

Paper IV

The sample for this paper consisted of all individuals in the HUNT-2 study, who after imputation had valid scores on all 10-SCL items. This resulted in a total of 46,064 individuals, 21,718 males and 24,385 females. The mean age in the sample was 48.5 years, (48.8 for males, 48.8 for females).

4.4 Measures

Personality disorders (Paper I and II)

DSM-IV Axis-II disorders were assessed using a Norwegian version of the Structured Interview for DSM-IV Personality (SIDP-IV) (Pfohl B., Blum N., & Zimmerman M., 1997). This instrument is a comprehensive semi-structured diagnostic interview for the assessment of all DSM-IV axis II disorders, including the two appendix diagnoses Depressive PD and Passive-Aggressive (Negativistic) PD. SIDP was initially developed in 1983, and has been used in a number of studies in many countries including Norway (Helgeland & Torgersen, 2004; Torgersen, Kringlen, & Cramer, 2001). The instrument includes non-pejorative questions organized into topical sections (e.g “social relationships”, “work style”, “emotions”) rather than disorders. This allows for a more natural flow of the interview and increases the likelihood that useful information from related questions may be taken into account when rating related criteria within that section. The specific DSM-IV criterion associated with each set of questions is rated according to the following scoring guidelines: 0 = “not present or limited to rare isolated examples”, 1 = “subthreshold – some evidence of the trait, but it is not sufficiently pervasive to consider the criterion present”, 2 = “present – criterion is clearly present for most of the last 5 years (i.e. present at least 50% of the time during the last 5 years), 3 = “strongly present – criterion is associated with subjective distress or some impairment in social or occupational functioning, or intimate relationships”. The SIDP-IV interview is conducted after the axis I interview in which axis I disorders are assessed. This helps the interviewer to more easily distinguish longstanding behavior reported by the subject from temporary states due to an episodic psychiatric disorder. The SIDP-IV uses the “five year rule” which means that the behavior, cognitions and feelings that have predominated for
most of the last 5 years are considered to be representative of the individual’s long-term personality functioning.

Phobias (Paper III)

DSM-IV Axis-I disorders were assessed using a computerized version of the Composite International Diagnostic Interview (CIDI). This computerized version (DIAX) yields all major ICD-10 and DSM-IV diagnoses. From these, specific phobias, agoraphobia and social phobia were selected. Due to the low prevalence of situational phobia, situational and environmental phobias were merged into a single variable which we referred to as environmental/situational. This left us with five phobias for analysis: animal phobia, environmental/situational phobia, blood phobia, agoraphobia (with and without panic), and social phobia. In addition to assigning full DSM-IV diagnoses, the computerized CIDI interview also assigns sub-threshold phobia diagnoses in cases where all but one of the criteria of the full disorder are met. In order to increase statistical power, sub-threshold scores were included in the twin analyses. The variables analysed were coded as 0 (no diagnosis), 1 (sub-threshold phobia diagnosis) and 2 (full phobia diagnosis).

Paper IV

The ten item symptom checklist (SCL-10) is a shortened version of the original 25 item symptom checklist (SCL-25) (Hesbacher et al., 1980) and is designed to measure the two dimensions of anxiety and depression in large health surveys. The test has demonstrated good psychometric properties in previous Norwegian studies, and has been shown to correlate highly (r=0.97) with SCL-25 (Strand, Dalgaard, Tambs, & Rognerud, 2003). The participants are asked to rate on a scale ranging from 1 to 4, how bothered or distressed they were the past 14 days by each of the ten symptoms, four of which address anxiety and six depression. For anxiety, these symptoms were; “Suddenly scared for no reason”, “Nervousness or shakiness inside”, “Faintness, dizziness, or weakness”, and “Feeling tense or keyed up”. The depression subset consisted of the items; “Blaming yourself for things”, “Difficulty falling asleep, staying asleep”, “Feeling blue”, “Feeling of worthlessness”, “Feeling everything is an effort”, and “Feeling hopeless about the future”.

Expectation maximization (EM) algorithm (Rubin, 1991) in SPSS 12.0.1 was used to impute values in cases with scores on 5 or less SCL-10 items missing. Imputation increased the total effective sample size from 42,184 to 46,064.
4.5 Statistical analyses

4.5.1 Liability-threshold model

The multifactorial etiology of complex phenotypes has important theoretical implications. If the expected value of a continuously scored measure of mental health is determined by a large number of genes or environmental factors individually contributing a small amount of increase in risk, then the distribution of the trait in the population will be approximately normal. While many researchers are advocating a dimensional understanding of mental disorders (Krueger, Watson, & Barlow, 2005; Widiger & Samuel, 2005), both the ICD-10 and DSM-IV diagnostic manuals represent mental disorders as discrete, non-overlapping categories. If diagnostic information is used, quantitative geneticists have to analyse scores which are coded on a binary or ordinal scale. This is true for all measures analysed in the present study, and adds a level of complexity to the statistical modelling, as the original analytic framework was developed for continuous and normally distributed data.

The simplest way of estimating heritability of categorical measures is by considering the relative difference in the concordance between MZ and DZ twins. However, most current analyses of threshold traits rely on the liability-threshold concept first proposed by Wright (Wright, 1934). In this approach, while the disorder itself is binary, an underlying gradation of some attribute, either genetic or environmental, immediately related to the disease is assumed, and referred to as an individual’s liability to the disease. It is further assumed that a measure of this attribute would give us a score with normal distribution, and that individuals above a certain threshold value would exhibit the disease, while those below would not. As discussed 1.2, the evidence that mental disorders are multifactorial supports the assumption that the underlying liability is continuous and with a standard normal distribution. This is because the critical assumption of a normally distributed liability follows directly from the central limit theorem, which states that the mean or sum of a sufficiently large number of independent random variables will be approximately normal (Rice, 1995).

4.5.2 Twin models

The basic univariate twin model has been described in detail in section 2.4. The models below describe extended versions of this model.

Univariate sex-limitation twin model (Paper I)
In this paper we fitted a univariate model with scalar and non-scalar sex limitation. In the scalar or quantitative sex-limitation models, the same genes are assumed to influence the phenotype in both males and females, but the relative magnitude of the phenotypic variance explained by these genes are allowed to differ across gender. Scalar sex-limitation is tested by allowing independent A, C and E parameters across gender, but fixing the expectation for the additive genetic unlike-sex DZ twins (DZU) correlation to 0.5. The non-scalar or qualitative sex limitation models tests whether different genes influence the variance in males versus females, and is implemented by letting the expected DZU correlation range from 0 to 0.5. Instead of gender differences in genetic influences, gender differences in the effect of shared environmental can be estimated. However, as these two models are not nested, and cannot be compared directly, we chose a genetic non-scalar sex-limitation model as our full reference model. The power to determine whether non-scalar sex-specific influences are present is entirely dependent on the information available in the DZU group, and is therefore often low.

While sex-limitation is readily implemented in univariate twin models, they are considerably more difficult to apply to multivariate models. For an in depth discussion of the problem as well as solutions under certain multivariate models, see Neale et. al. (2006) (Neale, Roysamb, & Jacobson, 2006).

Independent pathway model (Paper II and paper III)
The independent pathway (IP) model, is a multivariate twin model where parameters on all paths from common sources are free to vary independently of each others, constrained only by the total variance of each phenotype.

Common pathway model (Paper III)

Under the common pathway (CP), common genetic and environmental factors influence all observed variables through a single psychometric factor, or underlying latent liability (Rijssijk, 2005). This model constrains the pattern of influence of common A, C and E to be equal across the observed phenotypes. The CP model can therefore be parameterized as a series of constraints imposed on an independent pathway model, and can be formally compared in fit to the more general independent pathway model.

4.5.3 Nuclear family model (Paper IV)

A nuclear family data structure was constructed using information supplied by Statistics Norway, identifying first-degree relationships. Polychoric correlations between SCL-10 scores of family members were estimated by means of the “polycor” package in R (John Fox, 2008). By the rules of path analysis, correlations expected under the given model can be expressed as a set of nonlinear equations of model parameters. These parameters were estimated by weighted least squares (WLS) using the nonlinear minimization function in R, an open source software package for statistical computing (R Development Core Team, 2005). WLS typically give estimates that are close to that of maximum likelihood in kinship studies, while being far less computationally demanding (see section 4.5.4).

There is not enough information in the nuclear family data structure to differentiate between genetic effects and environmental effects transmitted from parents to offspring. We
therefore fitted a model where phenotypic SCL-10 variance was assumed to be solely a function of additive genetic effects (G) and unique environmental effects (E).

Any residual sibling similarity beyond what can be accounted for by genetic factors can be modelled as a “sibling effect”, though to reflect similarities in the environment of siblings. As for genetic effects, in the full model, sibling parameters were initially allowed to be sex-specific, giving potentially different correlations between brothers and sisters ($S_f / S_m$). Opposite-sex sibling effects were further moderated by a parameter $\sigma$, allowing for a potential difference of similarity between OS siblings.

4.5.4 Technical issues

Summary statistics vs. raw data

Until the previous decade, *summary statistics* such as correlations or covariances invariably constituted the data points to which parameter values in structural equation models were fitted. Pearson correlations are typically used for continuously and normally distributed data, while polychoric correlations are more appropriate for threshold traits. Polychoric correlations are estimated by fitting a bi or multivariate normal distribution to the frequency table of paired scores (Olsson, 1979).

Given the pre-calculated correlations, parameter values of the structural equations are estimated by selecting those values that minimize a measure of distance between the observed and expected correlations. The most common estimators when analysing summary statistics are weighted least squares (WLS) or maximum likelihood (ML). In WLS, parameters are estimated by minimizing the sum of squared differences between the observed correlations and those expected under a given model, multiplied by an appropriately chosen weight, usually one over the variance of the correlation. Unlike WLS, maximum likelihood begins with a parametric description of the model, and proceeds by varying the parameter values to find those that yield the largest joint likelihood. Without exception, data in twin and family models are assumed to be multivariate normal distributed.

Maximum likelihood models have a number of advantages over weighted least squares. That they are asymptotically unbiased, so that given a sufficiently large sample, the estimate will equal the true population value, and they are efficient, in that the estimates have a small variance. Importantly, likelihood-ratio tests may be used to compare different models, as twice the difference in log-likelihood is, under certain regularity conditions is asymptotically distributed as $\chi^2$ with degrees of freedom equal to the difference in the number of parameters.
As computers have become substantially more powerful, twin analyses are now almost exclusively based on raw-data rather than summary statistics. With little exception, estimation with raw-data is performed by full-information maximum likelihood (FIML). Again, we begin with a parametric description, which in twin analysis invariably involves assuming the data are multivariate normal. Parameter estimation is performed by maximizing the likelihood function, or conversely, minimizing the negative log likelihood. The likelihood function is the joint probability of each data vector under the structural model. If the observations are independent, the joint likelihood equals simply the sum of the individual probabilities. Estimation is performed by selecting the parameter values that maximizes the likelihood, or conversely minimizes the negative log likelihood. The use of raw data and FIML has several benefits. Most importantly, records with one or more missing values need not be discarded. FIML also permits a number of more sophisticated models to be fitted, such as those including covariates on the means or thresholds.

However, while raw-data and FIML to a large extent have replaced the use of summary statistics in the analyses quantitative genetic approach, there are situations that warrant the use of WLS. Firstly, when analysing threshold traits, the FIML approach is very computationally demanding, as the likelihood contribution of a given record cannot be directly calculated in closed form, and must instead be estimated by numerical integration over a high dimensional space. FIML is therefore to a much higher degree subject to what is known as the “curse of dimensionality”, and becomes computationally unfeasible for models involving many phenotypes or large pedigrees. FIML is also more sensitive to starting values, meaning that the probability of getting stuck in local minima during optimization is substantial, and since each estimation is a very intensive task, taking hours or days for large models, only a small number of different starting values can be tried.

Model selection

Statistical models with subtle differences can have vastly different substantive interpretations. Therefore, in order to inform theory, not only an overall goodness of fit of a full reference model is needed, but also statistically well reasoned criteria to select between related models.

If correlation or covariance matrices are used, an overall test of goodness, the chi-square test is available. For FIML, the absolute log-likelihood value (LL) for a given model is not in itself informative. However, two nested models can be compared by a chi-square test, as -2 times the log likelihood under regularity conditions asymptotically approximates a chi-
squared distribution with degrees of freedom equal to the difference in the number of parameters.

A related and popular fit statistic to the chi-square, that penalizes model complexity and encourages parsimony, is Akaike Information Criterion (AIC) (Akaike, 1987). Unlike chi-square, AIC can be used to compare models that are nested as well as those that are not. As a rule of the thumb, an AIC greater than 0 reflects poor fit (Neale et al., 2003). Simulation studies have shown that the AIC should be used with caution, and when power is low, estimates from the full model should be reported as the principal results. (Sullivan & Eaves, 2002).

Both likelihood ratio tests and the related AIC statistics continue to be widely used measures of model fit in quantitative genetics, also in the present study. However, recent results demonstrate that these statistics should not be used uncritically, as they have been shown to yield p-values that are too high in most twin models (Dominicus, Skrondal, Gjessing, Pedersen, & Palmgren, 2006).

4.5.5 Tests of statistical assumptions

Multiple threshold tests (Paper I, II, III)

The multiple threshold test is used to determine whether pairs of observed ordinal scores are consistent with an underlying bivariate normal density. If the scores have more than two ordered categories, the test will yield an estimate for the correlation in this underlying distribution, as well as a test of the goodness of fit of this model (Reich, James, & Morris, 1972). In paper I, II and III, multiple threshold tests were conducted to assess whether sub-threshold scores and threshold scores can be considered points on the same underlying liability distribution, and hence are quantitative rather than qualitatively different.

Tests of equal environment (Paper I and II)

The equal environment assumption (EEA) posits that MZ and DZ twins are equally correlated in their exposure to environmental events of etiologic importance for the trait under study (Kendler, Neale, Kessler, Heath, & Eaves, 1994). If this assumption is violated, a difference in observed similarity between MZ and DZ twins that is due to a different level of environmental influence can be wrongly attributed to genetic factors. To test for possible violation of the EEA in our sample, a polychotomous logistic regression was fitted,
controlling for the correlational structure of our data using independent estimating equations as operationalized in the SAS procedure GENMOD (SAS Consulting Department of Statistics, 2002). Two variables that reflected shared environment during childhood (number of years that the twins were in the same class at school and the years the twins lived in the same residence) and during adulthood (frequency of in-person and telephone contact during the last year and the distance between their current residences) were constructed. We then tested whether the dimensional PD score in twin 1 interacted with our measure of environmental similarity in predicting the dimensional PD score in twin 2 (dependent variable). We controlled for main effects of zygosity, sex, age and level of environmental similarity as well as shared environment effects and genetic effects.

Tests of the EEA was performed for all axis-II disorders, and found to be non-significant both for twin contact during childhood and contact during adulthood, indicating that EEA was not violated.

Tests of EEA was not carried out for the phobia subtypes, but have been done in several previous studies, and found to be negative (Kendler et al., 1994; Kendler, Neale, Kessler, Heath, & Eaves, 1993).
5. Main Findings

Paper I

Passive-aggressive personality disorder (PAPD) is one of the least studied PDs. The objective of this paper was to investigate the familial aggregation of PAPD, and explore issues regarding this disorder raised by the DSM-IV Personality Disorder Work Group. 2794 twins from the population-based Norwegian Institute of Public Health Twin Panel were interviewed with the Structured Interview for DSM-IV Personality (SIDP-IV). Because of the rarity of the twins meeting full diagnostic criteria for PAPD a dimensional representation of the disorder was used for the analyses. Overlap with other axis II disorders was assessed by polychoric correlations, while familial aggregation was investigated by structural equation twin models. Correlation in symptom count was highest with paranoid ($r=0.52$) and borderline personality disorder ($r=0.53$), and lowest with schizoid ($r=0.26$). Both MZ and DZ twin correlations were significantly non-zero, indicating familial aggregation of passive-aggressive symptoms. The twin correlations and parameter estimates in the full model showed genetic and shared environmental effects for females, and only shared environmental effects for males, but the prevalence of endorsed PAPD criteria in this community sample was too low to permit us to conclude with confidence regarding the relative influence of genetic and shared environmental factors on the familial aggregation of PAPD.

Paper II

The DSM-IV cluster C Axis II disorders include avoidant (AVPD), dependent (DEPD) and obsessive-compulsive (OCPD) PDs. The objective of this paper was to estimate the genetic and environmental influences on dimensional representations of these disorders and examine the validity of the cluster C construct by determining to what extent common genetic and environmental factors influence the individual PDs. PDs were assessed using the Structured Interview for DSM-IV Personality (SIDP-IV) in a sample of 1386 young adult twin pairs from the Norwegian Institute of Public Health Twin Panel (NIPHTP). A single-factor independent pathway multivariate model was fitted to the number of endorsed criteria for the three cluster C disorders, using the statistical modelling program Mx. The best-fitting model
included genetic and unique environmental factors only, and equated parameters for males and females. Heritability ranged from 27% to 35%. The proportion of genetic variance explained by a common factor was 83%, 48% and 15% respectively for AVPD, DEPD and OCPD. Common genetic and environmental factors accounted for 54% and 64%, respectively, of the variance in AVPD and DEPD but only 11% of the variance in OCPD. Cluster C PDs are moderately heritable. No evidence was found for shared environmental or sex effects. Common genetic and individual environmental factors account for a substantial proportion of the variance in AVPD and DEPD. However, OCPD appears to be largely etiologically distinct from the other two PDs. The results do not support the validity of the DSM-IV cluster C construct in its present form.

Paper III

The objective of this paper was to explore the genetic and environmental factors underlying the co-occurrence of lifetime diagnoses of DSM-IV phobia. Twins from the population-based Norwegian Institute of Public Health Twin Panel were assessed at personal interview for DSM-IV lifetime specific phobia (animal phobia, environmental phobia, situational phobia and blood phobia), social phobia and agoraphobia. Because earlier studies had indicated possible gender differences, and we had no power to investigate these effects in our sample, only female twins (n=1430) were included in the analyses. Comorbidity between the phobias were assessed by odds-ratios and polychoric correlations. Phenotypic correlations of lifetime phobia diagnoses ranged from 0.55 (agoraphobia and social phobia, OR=10.95) to 0.06 (animal phobia and social phobia, OR=1.21). In the best fitting twin model, which did not include shared environmental factors, heritability estimates for the phobias ranged from 0.43 to 0.63. Comorbidity between the phobias was accounted for by two common liability factors. The first loaded principally on animal phobia, but also weakly on the other specific phobias, but did not influence the complex phobias (agoraphobia and social phobia). The second liability factor strongly influenced the complex phobias, but also loaded weakly to moderately on all the other phobias. Our results therefore suggest that the comorbidity between phobias is best explained by two distinct liability factors rather than one single factor, as has been assumed in most previous multivariate twin analyses. Blood phobia was mainly influenced by a specific genetic factor which accounted for 51% of the total and 81% of the genetic variance.
Numerous studies have found self-reported symptoms of anxiety and depression to be heritable. However, these estimates are based almost exclusively on analyses of twins. The objective of this paper was to estimate an upper limit on the heritability of self reported symptoms of anxiety and depression in a large and population representative sample of nuclear families. The ten-item symptom checklist (SCL-10) was administered as part of a large health survey in a Norwegian county. The SCL-10 is a shortened version of the SCL-25, and is intended to measure the dimensions of anxiety and depression. In all, 46,064 people responded, and data from Statistics Norway allowed us to link responses of first degree relatives. Polychoric correlations were calculated, and a simple nuclear family model was fitted to these data using the software package R. All correlations between nuclear family members were in the range 0.12 to 0.16, suggesting small but significant familial influences on SCL-10. In the best fitting model, heritability was estimated at .25 (95% CI = .22-.27), and sibling specific environmental effects could be discarded. This heritability is lower than most published estimates of comparable measures in twin samples.
6. DISCUSSION

6.1 Methodological considerations

The results presented in this thesis should be regarded in the context of a number of limitations and methodological considerations. In this section I review different aspects of the design and statistical analyses that may impact the generalizability of our findings.

6.1.1 Reliability

In quantitative genetic analyses, the error variance stemming from an unreliable measure is included in the estimate of unique environment. This error therefore results in a smaller proportion of total variance attributable to additive genetic and shared environmental factors. To ensure that we attain correct estimates of these variance components, we must demonstrate that the measures have satisfactory reliability.

*Personality disorder measures*

A review of inter-rater reliability of axis-II diagnoses attained by semi-structured interviews found the kappas for the 12 DSM-III-R PDs across 15 studies to lie in the range 0.62-0.76 (Zimmerman, 1994). As a kappa coefficient greater than 0.6 by convention is considered to indicate substantial agreement, these results suggest that semi-structured interviews for PDs have considerable diagnostic reliability. In our twin sample, the number of subjects with specific PDs was too low to calculate Kappa coefficients. Instead, inter-rater reliability was assessed based on 2 raters scoring of 70 audio-taped interviews. Intra-class correlations for the scaled PDs were all high: 0.96 for AVPD, 0.96 for DEPD, and 0.92 for OCPD, 0.91 for PAPD.

Test-retest reliability was not assessed in this study, but investigations of other structured interviews of axis-II disorders typically conclude that they are adequate, with kappas typically found in the range 0.45-0.9 (Chess & Thomas, 1991). Generally, inter-rater reliability is substantially higher than test-retest reliability for structured interviews of personality disorders (Bronisch & Mombour, 1998).

Reliability measured by Cronbach’s alpha based on polychoric correlations showed good internal consistencies: PAPD, 0.77; AVPD, 0.96; DEPD, 0.94; OCPD, 0.90.
SCL-10.

SCL-25 was originally designed as ‘state’ measure, and respondents are asked to rate how bothered or distressed they were the past 14 days. However, a range of studies have demonstrated that these symptoms display considerable temporal stability, and to a large extent reflect stable or ‘trait’-like aspects (Foley, Neale, & Kendler, 2001).

In the HUNT-2 sample, our only measure of reliability is Cronbach’s alpha, which was found to be 0.85. This indicates good internal consistency, and is similar in magnitude to what has been reported in other Norwegian samples in both the SCL-10 and SCL-25 (Strand et al., 2003).

Phobia

The inter-rater reliability on CIDI has been demonstrated to be excellent (Andrews & Peters, 1998).

Test-retest reliabilities of lifetime diagnoses of simple phobia, social phobia, and agoraphobia over a period between 16 and 34 months assessed by the CIDI interview have elsewhere been found to be modest (Wittchen, Zhao, Abelson, Abelson, & Kessler, 1996; Kendler, Karkowski, & Prescott, 1999a).

6.1.2 Validity

Personality disorder measures.

Studies that have examined the validity of structured interviews for axis-II disorders by comparing scores on these instruments with consensus diagnosis from teams of psychiatric professionals, have found that validity of “any-PD” is satisfactory (κ = 0.55-0.58), but the validity of a specific diagnosis is modest (Bronisch & Mombour, 1998). Regardless, structured interviews remain the “gold standard” in assessment of axis-II disorders.

In paper I and II, we based our analyses on the truncated sum-scores of subclinical symptoms rather than threshold diagnoses. This rests on two assumptions; First, that subclinical symptoms are quantitatively rather than qualitatively different from clinical symptoms, and second, environmental and genetic determinants on the variance in a measure that is almost exclusively dominated by scores below the number which would qualify for a DSM-IV diagnosis is similar to those of clinical disorders.
To assess whether sub-threshold scores and threshold scores could be considered points along the same continuous underlying liability distribution, we performed multiple threshold tests the four PD variables. All the tests were performed separately for each zygosity group, and none was significant (all p values >0.05).

The SCL-10.

The 10 item symptom checklist (SCL-10) has demonstrated good psychometric properties in previous Norwegian studies, and has been shown to correlate highly (r=0.97) with SCL-25 (Strand et al., 2003). The SCL-25 in turn has proved to have satisfactory validity as a measure of psychological distress (Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974). The validity of the SCL anxiety and depression sub-scales has been empirically demonstrated in clinical investigations involving more than 2500 patients (Derogatis, Lipman, & Covi, 1973). Another study evaluating a threshold score on SCL-25 as a screening instrument for DSM-III-R axis-I disorders in young adults, found moderate sensitivity (43–70%), and high specificity for anxiety and depression (83–85%) (Veijola et al., 2003).

6.1.3 Statistical power

The biometric analyses in all papers included in this thesis have been conducted on ordinal measures of psychopathology. Simulation studies have shown that statistical power in twin studies is considerably lower in analyses of ordinal than continuous data, and that power is strongly and positively correlated with prevalence (Neale et al., 1994). While ordinal rather than binary measures have been employed to increase power, prevalence was low, particularly for many of the personality disorders. We can therefore not reject the possibility that a moderate effect of shared environment on the cluster C PDs or the phobias could have been left undetected.

6.1.4 Selected technical issues

The use of sum scores rather than measurement models

In paper I and II we chose to estimate variance components on a sum-score rather than a measurement model, i.e. a confirmatory factor model specifying the relationship between observed indicators and a latent variable (Schumacker & Lomax, ). This choice was entirely pragmatic, as the computational demands of FIML estimation scales steeply when the number of threshold phenotypes exceeds five. We must acknowledge that analyses of sum-scores,
despite being widely common, do introduce potential sources of error. Specifically, the absence of measurement invariance across zygo sity can bias the estimates of genetic and environmental components of variance (Neale MC, Lubke G, Aggen SH, & Dolan CV, 2006). Item level analyses would also bypass some of the measurement error included in a sum-score, and would in all likelihood result in higher estimates of heritability. Unfortunately, the endorsement of PD criteria was too low, and the computational resources required too high, to conduct multivariate analyses at the individual item level.

Choices forced by the use of FIML estimation

In retrospect, I believe polychoric correlations rather than raw-data could be used in paper I and II. The reason is that FIML is sensitive to missing values in the frequency tables of paired scores, and to avoid such missing values, we had to truncate the sum-scores. However, as we had virtually no missing responders, the benefits of FIML estimation does not in my mind outweigh the bias that in theory could be introduced by such truncation.

Counting subthreshold criteria rather than summing the SIDP scores

Individual PD criteria are in the SIDP interview scored on a scale of 0 to 3, with the values 1 representing a sub-threshold score, and the values 2 and 3 threshold scores. In our analyses we decided to let each criteria with a nonzero score carry equal weight. While we recognize that this may be seem wasteful from a psychometric perspective, due to the rarity of scores of 2 or 3, summing the scores rather than counting the number of nonzero criteria gave virtually identical results.

6.1.5 Clinical samples vs. population based samples.

One of the greatest strengths of the present study is that it is based on two population based samples. Numerous studies have sought to determine whether samples drawn from clinical settings are representative of the population as a whole. They suggest that participants in clinical samples are more impaired, and more likely to have comorbid disorders (Goodman et al., 1997; Dufort, NEWMAN, & Blandr, 1993). Consequently, findings from individuals seeking treatment risk overestimating comorbid pathology. This is known as Berkson’s Bias, and thought to stem from the fact that people with multiple diagnoses are more likely to seek treatment than those with only a single diagnosis (Schwartzbaum, Ahlbom, & Feychting, 2003).
6.1.6 Assumptions of twin and family analyses

All mathematical models are abstractions, and can only to a limited degree represent the complexities of in the phenomena they model. In the physical sciences as well as the social ones, and in particular when modelling a process as complex as human development, certain simplifications must be made. These assumptions are clearly stated in the statistical models, and their validity can be scrutinized. In this section I briefly review some of the assumptions, and consider their validity with respect to psychopathology.

The assumption of no gene x environment interaction

GxE interaction refers to a genetically predisposed sensitivity to certain environmental experiences. In psychopathology, such an interaction is implied by the diathesis-stress model, which posits that both an inherited disposition and an environmental stressor is necessary to develop a given disorder, and neither one is sufficient in itself (Zubin & Spring, 1977). The diathesis-stress model has been used to explain the genetic influences, but incomplete penetrance of almost every major psychiatric disorder (Hammen, Henry, & Daley, 2000; Watson, 1999). GxE interaction was for a long time dismissed by behaviour geneticists, on the grounds that they were sufficiently rare and unimportant to be safely disregarded (Rutter, 2006). However, there is now growing evidence from both experimental studies and molecular genetics that such interaction effects are commonplace (Rutter M, Mpoftitt T, & Caspi A, 2006). Indeed, GxE interaction are thought to arise in all domains of disease, including cancer, cardiovascular disease and auto-immune disorders (Susser, 2006). A widely cited example pertaining to personality pathology is a study by Caspi et. al. where a functional polymorphism in the gene encoding the neurotransmitter-metabolizing enzyme monoamine oxidase A (MAOA) was found to moderate the effect of maltreatment on subsequent antisocial behaviour (Caspi et al., 2002). Results are intriguing, but have not been consistently replicated (Haberstick et al., 2005). GxE interaction in a measure of psychopathology has also been implicated in a long term follow up of Finnish adoptees, where a significant effect of disordered adoptive rearing on development of a schizophrenia-spectrum diagnosis was found only in the adoptees at high genetic risk (Tienari et al., 2004). In summary, the heterogeneity in response to all causal factors of mental disorders suggests there is reason to expect GxE interaction to be prevalent, and relevant also to the disorders analysed in this thesis. Twin models can to some extent accommodate interaction effects, but they naturally require a measure of a relevant environmental exposure.
Failure to take GxE interaction effects into consideration may deflate estimates of additive genetic effects, while GxC interaction will be estimated as A (Rijlsdijk & Sham, 2002).

The assumption of zero gene-environment correlation (rGE)

As was detailed above, all covariance terms in equation 1 section 1.4 p. 5 were assumed to be zero. Gene environment correlations are traditionally referred to as either passive or active. Active gene-environment correlation refers to the fact that the genetic disposition of individuals may impact the probability that they select themselves into certain environments. Passive gene-environment correlation refers to the fact that the developmental environment into which people are born is provided by their biological relatives, to whom they are also genetically related.

Twin studies have found convincing evidence for rGE in mental disorders, by demonstrating that the aversive life events that have been associated with these disorders are themselves moderately heritable (Bolinskey, Neale, Jacobson, Prescott, & Kendler, 2004).

Ignoring an active and positive AxE correlation will inflate heritability estimates, while a negative correlation will decrease the estimate. Ignoring a positive passive CxE correlation will inflate shared environmental effects (Rijlsdijk & Sham, 2002).

Equal environment assumption

The equal environment assumption (EEA) posits that MZ and DZ twins are equally correlated in their exposure to environmental events of etiologic importance for the trait under study (Kendler et al., 1994). The EEA has probably been the most widely discussed assumption of the twin model. If this assumption is indeed false, and MZ twins are more similar in their environmental exposure, then this similarity will be wrongly attributed to genetic factors, and genetic effect is inflated at the expense of effect of family environment. The validity of the EEA cannot be determined once and for all for every phenotype, nor is there a single way to operationalize what constitute a genuine test of the assumption. Instead, a number of techniques have been employed across a variety of phenotypes to gauge the probability that EEA may bias the results, and they jointly supported the validity of the EEA for psychiatric disorders (Kendler, Neale, Kessler, Heath, & Eaves, 1992a); (Kendler et al., 1993; Kendler & GARDNER JR, 1998); (Kendler et al., 1994). The consensus among researchers in the field is that violation of EEA is unlikely to be of sufficient magnitude to pose a threat to the twin strategy for studies of psychiatric disorders (Rutter, 2006). As
detailed in section 4.5, we tested the EEA for personality disorders, and found no violation of the assumption. Several studies have examined the EEA with respect to phobia, depression, generalized anxiety, and found the assumption hold (Kendler et al., 1994; Kendler et al., 1993).

*The assumption that gene action is additive*

Only additive genetic effects are included in the models in this thesis, and this is true for much published twin research. It is often taken for granted in behaviour genetics that the genes underlying a trait are individually Mendelian, and interact in a simple way to create a gauss distribution for the trait in the population. In other words, we preclude a large effect of gene*gene interaction. There is sound biological reason to expect hereditary influences on many traits to be additive. For a trait to be subject to evolutionary forces, we need gradual rather than erratic changes over generations. However, if epistatic influences were the norm, this would make evolutionary selection difficult. Theoretically, epistatic effects could be modelled, though in practice this is almost never done. This is in large part due to the weaker theoretical basis for these models, and the fact that phenotypic correlations between DZ twins are usually somewhat larger than half that of their MZ counterparts, which do not suggest any widespread epistatic effects. Still, there is uncertainty as to how reasonable the assumption of no epistasis is. Furthermore, the plausibility of this assumption probably depends on the given trait. Williams et al. argue that the genetic studies of complex traits in animals have had more consistent results, largely because one is more able to control for genetic background in animal studies. If there were epistatic effects involved, then an inability to control genetic background would lead to low replicability, which is what has characterized the search for genes for complex diseases in humans (Williams, Haines, & Moore, 2004).

Dominance is typically only included in the twin models if DZ correlations are substantially less than half of MZ correlations. The only phenotypes in this thesis that displayed this pattern were animal phobia and to a lesser extent social phobia. To my knowledge, only one paper has demonstrated significant dominance on a phobia-like phenotype (Kendler et al., 1995b). However, high statistical power is required to differentiate between A and D, and A,C and D cannot all be included in the same model. Because of limited power, and since C is expected based on a popular psychological theory for phobia development (social learning theory), we chose not to include D in the multivariate model.
We assume that zygosity and biological family relatedness is measured without error

Estimates of heritability can be wrong if there is substantial misclassification of genetic relatedness. However, the questionnaire based zygosity test and DNA markers indicate an expected misclassification rate of only 0.67%, too small to significantly impact the results (Neale, 2003a).

In the HUNT-2 samples, only mother-offspring were linked with complete certainty, and we expect a certain misclassification of fathers and siblings. To assess the impact of father-offspring misclassification, we calculated the correlations for height, known to be almost entirely due to additive genetic influences. Mothers correlated 0.47 (N= 18,589) with their sons and 0.45 (N=16,345) with their daughters, while fathers correlated 0.44 (N=15,596) with their sons and 0.43 (N=13,741) with their daughters. Although father-offspring is marginally lower than mother-offspring, we believe this is not enough to seriously bias the results. Sisters correlated 0.49 (N=5395), brothers 0.50 (N=7151), and OS-sibs 0.47 (N=12,042). These correlations indicate that there is little error in sibling classification.

The assumption of random mating

Assortative mating is the tendency for organisms to mate with individuals that are similar to themselves in some way. For instance, people tend to select partners that have a higher than average degree of similarity to themselves on measures of personality and interests (Botwin, Buss, & Shackelford, 1997). If the traits that are selected for are genetically influenced, partners will be genetically correlated with respect to these traits. This correlation between parents can be important to consider when estimating heritability, as many models, including the basic twin model, assume parents are genetically uncorrelated. If the parents’ genes were correlated, DZ twins would share more than 50%. In twins reared together this will artificially inflate the estimates of the shared environment and deflate the genetic component, as DZ twins are more similar than a genetic model based on random mating can account for.

Depending on the magnitude and cause, this correlation will give rise to different expectations of sibling and parent-offspring similarity. A significant spouse correlation may be caused by phenotypic homogamy, where spouses choose each other based on the trait under study, or as social homogamy, where spouse similarity results from phenotypic similarity within social groups.

Significant but moderate primary assortment have been found both within and across psychiatric diagnoses. (Maes et al., 1998). Assortative mating can be especially problematic if
psychopathology in one parent is associated with a different psychopathology in the other. This may make it seem as if two traits have a shared genetic liability when this is not the case (Rutter, 2006).

In papers I, II and III, we have no data on parents, and could not incorporate assortative mating. In paper IV, we modelled spouse correlation as phenotypic homogamy. Spouse correlation was modest, so the covariance between siblings that can accounted for by assortative mating in SCL-10 was very small. It is therefore likely that the effect of assortative mating on these traits would only to a small degree affect our estimates.

6.1.7 Limitations

Representativity of the sample

Participation in epidemiological studies is non-random, and individuals with lower educational and income levels are under-represented as are those living in institutional settings (Fowler, 2008). Furthermore, socially unacceptable traits, such as those of personality disorders, may therefore be underreported, affecting mean values and estimates of prevalence.

Quantitative genetics is less concerned with the mean level of a trait, but is instead based on the degree of phenotypic similarity between individuals of different genetic relatedness. However, these higher order statistics are also vulnerable to bias. If the variable being studied are correlated with those which underlie non-participation, estimates of twin-pair resemblance can be significantly altered (Neale, Eaves, Kendler, & Hewitt, 1989). Differential attrition and non-response may potentially lead to biased estimates of genetic and environmental parameters (Heath, Madden, & Martin, 1998). Twin studies may also be affected by different magnitudes of selection bias in monozygotic (MZ) versus dizygotic (DZ) pairs.

Tambs et. al. have investigated predictors for cooperation and attrition in the Axis-I axis-II mental health study. They found monozygosity, female sex, being unmarried, having no children, and high education to predict participation, whereas few indicators of poor mental and somatic health and unhealthy lifestyle moderately predicted nonparticipation in the second questionnaire study (Q2). Standard genetic twin analyses of indicators of various mental disorders from Q2, validated by diagnostic data from the AI/AII, did not indicate differences in genetic/environmental covariance structures between participants and nonparticipants in AI/AII. In general the results show a moderate selection towards good mental and somatic health. Attrition from Q2 to the AI/AII does not appear to affect twin analyses of mental health related variables (Tambs K et al., 2009).
6.2 Interpretation and conclusion

**Genetic epidemiology of PAPD and Cluster C PDs**

Numerous studies have found the higher order constructs of normal personality measures, such as the big five (Widiger & Costa, 2002), and personality disorders to be overlapping (Trull, 2005; Miller, Lyman, Widiger, & Leukefeld, 2001). Although there is disagreement to what extent the instruments for normal personality can capture the variance in personality pathology, some studies conclude that at the facet level, NEO-PIR (Costa & McCrae, 1992) can distinguish between many DSM-III-R axis II disorders (Widiger & Costa, 2002). Normal personality traits have in turn been found to be moderately heritable (Jang, Livesley, & Vemon, 1996). While this serves as an indirect line of evidence for the heritability of personality disorders, few quantitative genetic studies have been performed on axis-II disorders. The first twin analyses of PDs assessed by structured interviews were published by Torgersen et. al (Torgersen et al., 2000). However, they analysed data from a mixed clinical sample where at least one twin had been treated for a major mental disorder, and this could potentially seriously bias the results. Consequently, paper I and II are the first twin studies of passive-aggressive and cluster C personality disorders respectively, assessed by structured interviews in a population based sample.

In our analyses of dimensional PAPD we found the trait to be familial, as sibling correlations were moderate and statistically significant in all twin groups. However, contrary to the assumptions of twin models, male DZ twins correlated higher than MZ twins. We believe the high male DZ correlation was due to chance, since this group was considerably smaller than all the others, and the confidence intervals for the correlation were highly overlapping with those of the male MZ group. Unfortunately, the high male DZ correlations make it difficult to conclude as to what the cause of this familiality is. Indeed, the twin correlations alone suggest that C is the cause of male sibling resemblance, while female sibling resemblance was largely accounted for by A. We are however inclined to place most confidence in a model without sex-specific effects, and where A and C in roughly equal proportions account for twin similarity in PAPD traits.

In the cluster C PDs, a common genetic factor accounted for most of the genetic variance in AVPD and about half of the genetic variance in DEPD, consistent with findings in the only family study of cluster C PDs that showed a close familial relationship between
AVPD and DEPD (REICH, 1989). Our results indicate that for AVPD, DEPD and OCPD, familiality is best explained by genetic factors alone, with moderate genetic influence on all three PDs. No evidence was found for any sex differences in genetic and environmental influences on cluster C PDs. Evidence has however been found for quantitative sex differences in the heritability of Neuroticism (Lake et al. 2000), a trait closely related to AVPD and DEPD (Dyce & O’Connor, 1998).

These results are in line with the emerging understanding that a limited number of factors give rise to mental disorders, and that genetic and environmental determinants do not specifically increase the liability to a single disorder (Kendler et al., 1995a). In addition to indicating that our constructs are not identifying unique genetic liabilities, the results suggest that some of the same individual environmental experiences influence different PD traits, that is, environmental effects are not specific. Common unique environmental factors accounted for most of the environmental variance in DEPD and more than one-third of the environmental variance in AVPD. However, similar to the pattern found for genetic effects, unique environmental factors influencing OCPD were mostly specific to this PD.

**Issues pertaining to the personality disorders and DSM-V**

We believe the following aspects of the analyses of PDs are relevant to the next revision of DSM;

The first issue regards whether passive aggressive (negativistic) personality disorder should be included in DSM-V. Determining the relevant categories for nosology is an exceedingly challenging task, and a large number of criteria have been proposed for this purpose (Kendler, 1990). As this study was not primarily intended to investigate the diagnostic structure of the DSM, we do not have the data necessary to evaluate whether PAPD meets these criteria. However, one of the criteria proposed for psychiatric nosology is a significant familiality (Robins & Guze, 1970). Furthermore, while we found PAPD to be comorbid with the other axis-II disorders, it was not comorbid to a greater degree, and our results do not indicate that it can be subsumed into any of the other PDs, as has been suggested (Fossati et al., 2000). These findings jointly strengthen PAPD as a valid diagnostic category.

The second issue regards the cluster C personality disorders. Our results do not provide support for the validity of the DSM-IV cluster C construct in its present form. Given that several phenotypic studies indicate that OCPD stands apart from the other DSM clusters
(Kass, Skodol, Charles, Spitzer, & Williams, 1985; Hyler & Lyons, ; Hyler & Lyons, 1988; Sanislow et al., 2002) and twin studies of personality traits show that the phenotypic structure closely reflects the underlying genetic structure (Livesley, Jang, & Vernon P.A., 1998; Krueger, 2000; McRae, Jang, Livesley, Riemann, & Angleitner, 2001), our results can be viewed as supporting the hypothesis that OCPD represents a separate Axis II secondary domain.

Genetic epidemiology of phobias

Specific phobias are of particular interest to behaviour geneticists, as there is evidence for a genetic influence in the classes of stimuli that are feared by phobics. For the specific phobias, these are the classes that have posed a threat to man through our evolutionary history, and rarely those that pose the greatest risk today. The hypothesis that fear can more easily be associated with certain classes of stimuli is known as the preparedness hypothesis of phobia (Seligman, 1972).

Because a certain degree of fear of typical phobic stimuli may be beneficial, it is easy to imagine that there may have been a certain amount of evolutionary pressure on these traits. In fact, because both extreme fear and extreme lack of fear may sub-optimal, evolution may select for a moderate amount. It is therefore likely that the population mean on the trait is subject to what is referred to as stabilizing selection, which would lead to considerable additive genetic influences (Kendler & Prescott, 2006).

Different patterns of twin correlations would be expected by the major psychological theories on the etiology of phobias. If phobia is influenced by classical and operant conditioning, then it should be observable in twin data as E. Conversely, an alternative learning theory for the etiology of phobia posits that phobias are acquired by social learning, by seeing someone behave fearfully in the presence of a phobic situation (Rachman, 1977). If phobias are largely determined by social learning, then children are likely to be more correlated in their exposure to such a phobic role model, and we expect a larger amount of C.

We draw three main conclusions from the multivariate twin analyses;

First, all phobias were moderately to highly heritable, and in the most parsimonious sub-model, all C could be discarded. This is in line with evidence suggesting phobias aggregate in families (Fyer, Mannuzza, Chapman, Martin, & Klein, 1995; Stein et al., 1998), and the results of twin studies largely suggest that this familiality is due to a moderate additive genetic effect, accounting for 20% to 40% of the variance (Hettema JM et al., 2001; Merikangas & Low, 2005). Although there was some indication of shared environmental
influences in the full model, this could be discarded with hardly any loss in fit. This is in agreement previous twin studies on phobia (Kendler, Neale, Kessler, Heath, & Eaves, 1992b; Kendler, Karkowski, & Prescott, 1999b).

Second, the pattern of co-occurrence could best be accounted for by two common liability factors, both of which were highly heritable. The first liability factor accounted for nearly all the variance in animal phobia, but also had modest loadings on the two other specific phobias. It influenced neither social phobia nor agoraphobia. The second common liability factor loaded most heavily on the complex phobias, but had modest loading also on situational/environmental and blood phobia, and a weak influence on animal phobia. This suggests that the genetic risk factors underlying animal phobia are, to a large extent, disorder specific. Although our second common factor loads most heavily on agoraphobia and social phobia, it also accounts for comorbidity between the specific and complex phobias. The common risk factors are therefore more general than what was found by Hettema et al., who in a multivariate twin study of anxiety disorders found virtually no overlap between the genetic factors underlying the simple and complex phobias (Hettema, Prescott, Myers, Neale, & Kendler, 2005).

Third, not all the genetic influences on phobias could be explained by the two common factors. While specific genetic influences were estimated for both environmental/situational phobia and social phobia, by far the strongest indication of a specific genetic influence was found for blood phobia, where more than 80% of the genetic variance was disorder specific. This is noticeable, given that blood phobia distinguishes itself so strongly from the other phobias in the physiological response it elicits. Unlike the other phobias, blood phobia is characterized by a sudden drop in heart rate and blood pressure, often resulting in fainting (Page, 1994). However, while specific genetic influences on blood phobia may be plausible, very little specific genetic effect were found in a previous multivariate analysis of phobia in a large sample of male twins (Kendler, Myers, Prescott, & Neale, 2001).

**Family analysis of symptoms of anxiety and depression**

The main finding in the family analyses of SCL-10 was a significant familial aggregation, consistent with an upper limit on heritability of 0.25. This is somewhat lower than most estimates based on twin studies of comparable self-report scores, which typically report that the proportion of variance that can be attributed to genetic factors is lies in the range 30%-50% (Kendler, Heath, Martin, & Eaves, 1986; Jardine, Martin, Henderson, & Rao,
While our estimate is in the lower tail of the distribution, heritability of anxiety and depression in the low 0.3’s has been reported in several large twin studies (Kendler K.S. et al., 1994) (Agrawal et al., 2004).

It has been stated that there is a general tendency for family and adoption studies to yield lower estimates of genetic influence than twin studies (Loehlin, 1992; John et al., 2008), but results from two large population based family studies carried out in Norway are mixed. In a previous wave of data collected as part of the HUNT study in 1984-86, upper limit for the heritability of a set of items validated against SCL-25 was estimated at 0.22 (Tambs & Moum, 1993), while a second study estimated the upper limit heritability to be 0.43 (Tambs K., 1991).

If there is a tendency for family studies to yield lower heritabilities, it could be due to a number of reasons. First, the frequent assumption of purely additive genetic influences could be false. This is consistent with a review of all extended twin studies on neuroticism, a trait that correlates highly with measures of self-reported symptoms of anxiety and depression (del Barrio, Moreno-Rosset, López-Martínez, & Olmedo, 1997; Luteijn & Bouman, 1988), where evidence has been found for widespread non-additive genetic variation (Keller, Coventry, Heath, & Martin, 2005).

Second, differences between twin and family studies could be due to age specific genetic effects. This is not a problem in twin studies, as both twins have the exact same age. Although we cannot rule out age related genetic effects entirely, we compared the phenotypic similarity between siblings stratified by age difference, and found that age specific genetic effects are highly unlikely to be the sole reason for the low heritability.

Third, including an assortative mating parameter in our model accounts for some of the covariance between siblings that would otherwise be attributed additive genetic effects. We chose to model spousal similarity as an expression of phenotypic homogamy, which reduces the estimated heritability slightly since the genotypes of parents are allowed to be correlated. However, spousal correlation in our sample is modest (r=0.16), and a social homogamy model would not be enough to bring the results in line with twin studies.

In conclusion, heritability of SCL-10 in our sample is lower than most published estimates. We believe the field would benefit from more heritability estimates from family and extended twin studies, as to this would allow us to determine whether these approaches truly yield different estimates to twin studies.


Costa, P. T. & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI)*. Psychological Assessment Resources.


John Fox (2008). polycor: Polychoric and Polyserial Correlations (Version R package version 0.7-6) [Computer software].


Wright, S. (1934). The results of crosses between inbred strains of guinea pigs, differing in number of digits. *Genetics, 19*, 537-551.


doi:10.1017/S0033291706009548
http://journals.cambridge.org/action/displayAbstract?aid=987348

© 2006 Cambridge University Press
Access to the published version may require journal subscription.
Genetic and environmental influences on dimensional representations of DSM-IV cluster C personality disorders: a population-based multivariate twin study

TED REICHBORN-KJENNERUD1,2*, NIKOLAI CZAJKOWSKI1, MICHAEL C. NEALE4, RAGNHILD E. ØRSTAVIK1, SVENN TORGERSSEN3, KRISTIAN TAMBS1, ESPEN RØYSAM1,4, JENNIFER R. HARRIS1 AND KENNETH S. KENDLER4

1 Division of Mental Health, Norwegian Institute of Public Health, Oslo, Norway; 2 Institute of Psychiatry, University of Oslo, Oslo, Norway; 3 Institute of Psychology, University of Oslo, Oslo, Norway; 4 The Virginia Institute for Psychiatric and Behavioral Genetics and Departments of Psychiatry and Human Genetics, Medical College of Virginia/Virginia Commonwealth University, Richmond, VA, USA

ABSTRACT

Background. The DSM-IV cluster C Axis II disorders include avoidant (AVPD), dependent (DEPD) and obsessive-compulsive (OCPD) personality disorders. We aimed to estimate the genetic and environmental influences on dimensional representations of these disorders and examine the validity of the cluster C construct by determining to what extent common familial factors influence the individual PDs.

Method. PDs were assessed using the Structured Interview for DSM-IV Personality (SIDP-IV) in a sample of 1386 young adult twin pairs from the Norwegian Institute of Public Health Twin Panel (NIPHTP). A single-factor independent pathway multivariate model was applied to the number of endorsed criteria for the three cluster C disorders, using the statistical modeling program Mx.

Results. The best-fitting model included genetic and unique environmental factors only, and equated parameters for males and females. Heritability ranged from 27% to 35%. The proportion of genetic variance explained by a common factor was 83, 48 and 15% respectively for AVPD, DEPD and OCPD. Common genetic and environmental factors accounted for 54% and 64% respectively of the variance in AVPD and DEPD but only 11% of the variance in OCPD.

Conclusion. Cluster C PDs are moderately heritable. No evidence was found for shared environmental or sex effects. Common genetic and individual environmental factors account for a substantial proportion of the variance in AVPD and DEPD. However, OCPD appears to be largely etiologically distinct from the other two PDs. The results do not support the validity of the DSM-IV cluster C construct in its present form.

INTRODUCTION

The DSM-IV (APA, 1994) includes 10 personality disorders (PDs), coded on Axis II and grouped into three clusters A, B and C, often called the ‘Odd’, ‘Dramatic’ and ‘Anxious’. Cluster C consists of the avoidant (AVPD), dependent (DEPD) and obsessive-compulsive (OCPD) personality disorders.

Although numerous studies have examined the genetic epidemiology of DSM Axis I anxiety disorders (for a review, see Hettema et al. 2001), few have applied these methods to Axis II anxious PDs. Only one family study of cluster C
PDs has been published (Reich, 1989). The results indicated significant familiality for the DSM-III anxious cluster PDs as a whole, and for AVPD and DEPD (OCPD was not investigated separately). In a twin study of DSM-III-R PDs based on patient populations, Torgersen et al. (2000) found that the best-fitting models of cluster C PDs all included genetic factors. No population-based twin study of DSM PDs has been published.

The DSM-IV manual emphasizes that the classification of PDs into these three clusters is based on ‘descriptive similarities’ and has ‘serious limitations and has not been consistently validated’ (APA, 1994). Examination of the phenotypic structure, that is the pattern of covariance or co-occurrence between the disorders, has been used to test the justification of the cluster constructs. Although the results are equivocal, several studies have suggested that OCPD is only weakly related to the three traditional clusters (Kass et al. 1985; Hyler & Lyons, 1988; Nestadt et al. 1994; O’Connor & Dyce, 1998; Sanislow et al. 2002). Following Robins and Guze (1970), a more powerful way to validate the cluster C construct would be to determine the degree to which AVPD, DEPD and OCPD share familial/genetic risk factors. Multivariate twin studies have been used to evaluate the genetic structure of PD traits (Livesley et al. 1998) and normal personality (Krueger, 2000; McCrae et al. 2001), and this approach is among the research strategies expected to play a useful role in generating an empirical data base for the next edition of the DSM Axis II classification (Livesley, 2005; Widiger et al. 2005).

In this study we conducted multivariate twin analyses of a population-based sample of young adult twins to examine the genetic and environmental influences on dimensional representations of DSM-IV cluster C PDs, in order to examine the validity of the cluster C construct. We attempted to address two main questions: (1) What is the relative influence of genetic and environmental factors on AVPD, DEPD and OCPD in males and females? (2) To what extent are cluster C PDs influenced by common genetic, shared environmental and individual-specific environmental factors and to what extent are these factors specific to each individual PD?

METHOD
Sample
Data for these analyses come from the Norwegian Institute of Public Health Twin Panel (NIPHTP). The twins are identified through information contained in the Norwegian Medical Birth Registry, established 1 January 1967, which receives mandatory notification of all births. The current panel includes information on 15 370 like- and unlike-sexed twins born from 1967 to 1979. Two questionnaire studies have been conducted in this sample: in 1992 (twins born 1967–1974) and in 1998 (twins born 1967–1979). Altogether, 12 700 twins received the second questionnaire, and 8045 responded after one reminder (response rate 63%). The sample included 3334 pairs and 1377 single responders. The NIPHTP is described in detail elsewhere (Harris et al. 2002).

Data for the present report derive from an interview study of Axis I and Axis II PDs, which began in 1999. Participants were recruited among 3153 complete pairs who, in the second questionnaire, agreed to participate in the interview study, and 68 pairs who were drawn directly from NIPHTP. Of these 3221 eligible pairs, 0.8% were unwilling or unable to participate, and in 16.2% of pairs only one twin agreed to the interview. After two contacts requesting participation, 38.2% did not respond. A total of 2794 twins (44% of those eligible) were interviewed for the assessment of PDs. Attrition was not associated with measures of psychopathology (see Discussion). In 22 pairs where both twins initially agreed to be interviewed, one of the twins was later unable or unwilling to participate in the interview. The mean age of participants was 28.2 years (range 19–36 years).

Zygosity was initially determined by questionnaire items previously shown to categorize correctly 97.5% of pairs (Harris et al. 2002). In all but 385 like-sexed pairs, where one or both of the twins was either unwilling or unable to donate a blood sample, zygosity was also determined by molecular methods based on the genotyping of 24 microsatellite markers. Seventeen of these pairs with DNA information (2.5%) were found to be misclassified by the questionnaire data and were corrected. From the corrected data we estimated that, in our
entire sample, the zygosity misclassification rate was 0.7%, which is unlikely to substantially bias results (Neale, 2003). Our final sample consisted of 1022 males and 1772 females: 221 monozygotic male (MZM) pairs, 116 dizygotic male (DZM) pairs, 448 monozygotic female (MZF) pairs, 261 dizygotic female (DZF) pairs, 340 dizygotic opposite sex (DZO) pairs and 22 single responders.

Measurements

A Norwegian version of the Structured Interview for DSM-IV Personality (SIDP-IV; Pfohl et al. 1995) was used to assess PDs. This instrument is a comprehensive semi-structured diagnostic interview for the assessment of all DSM-IV Axis II diagnoses. The SIDP was initially developed in 1983, and has been used in a number of studies in many countries including Norway (Torgersen et al. 2001; Helgeland et al. 2005). The instrument includes non-pejorative questions organized into topical sections rather than by disorders. This allows for a more natural flow of the interview and increases the likelihood that useful information from related questions will be taken into account when rating related criteria within that section. The specific DSM-IV criterion associated with each set of questions is rated using the following scoring guidelines: 0 = not present or limited to rare isolated examples, 1 = subthreshold (some evidence of the trait, but not sufficiently pervasive to consider the criterion present), 2 = present (criterion is clearly present for most of the past 5 years, i.e. present at least 50% of the time), 3 = strongly present. The SIDP-IV interview is conducted after the Axis I interview, which helps the interviewer to distinguish more easily longstanding behavior reported by the subject from temporary states due to an episodic psychiatric disorder. The SIDP-IV uses the ‘5-year rule’, meaning that the behavior, cognitions and feelings predominating for most of the past 5 years are considered to be representative of the individual’s long-term personality functioning.

Interviewers were mostly psychology students in their final part of training and experienced psychiatric nurses. They were trained by professionals (one psychiatrist and two psychologists) with extensive previous experience with the instrument, and closely followed up individually during the whole data collection period. The interviews were carried out between June 1999 and May 2004, and were largely conducted face-to-face. For practical reasons, 231 interviews (8.3%) were obtained by telephone. Each twin in a pair was interviewed by different interviewers.

Inter-rater reliability was assessed based on two raters scoring 70 audiotaped interviews. The number of subjects with specific PDs was too low to calculate \( \kappa \) coefficients. Intra-class and polychoric correlations for the scaled PDs, using the number of endorsed criteria at the subthreshold level (see below), were all high: AVPD, +0.96, +0.97; DEPD, +0.96, +0.99; OCPD, +0.92, +0.87. Reliability measured by Cronbach’s \( \alpha \) based on polychoric correlations showed good internal consistencies: AVPD, 0.96; DEPD, 0.94; OCPD, 0.90.

Approval was received from the Norwegian Data Inspectorate and the Regional Ethical Committee, and written informed consent was obtained from all participants after complete description of the study.

Statistical methods

In this population-based sample of twins, the prevalence rate for categorical diagnoses of the cluster C PDs were too low for useful analyses. We therefore used a dimensional approach (Widiger & Samuel, 2005), constructing ordinal variables based on the number of criteria endorsed. To optimize statistical power and produce maximally stable results, we used the number of subthreshold criteria endorsed (\( \geq 1 \)) instead of criteria above the threshold (\( \geq 2 \)), assuming that the liability for each trait is continuous and normally distributed, that is that the classification (0–3) represents different degrees of severity. This assumption was evaluated using multiple threshold tests for each of the criteria. All the tests were performed separately for each zygosity group, and none was significant (all \( p \) values >0.05). Few subjects endorsed a high proportion of all or most of the criteria for an individual PD. To avoid empty cells, we collapsed the upper categories for the summed score. The maximum number of categories was created for each disorder. However, low prevalence in general for the DEPD items and low endorsement for the AVPD items in the DZM group limited the number of categories for these
two PDs to four (0–3). The ordinal variables for OCPD included five categories (0–4). The same procedure as described above was used to test the assumption that the number of positive criteria for each individual PD represented different degrees of severity for the PD. None of the multiple threshold tests was significant (all p values > 0.05).

In the classical twin model, individual differences in liability are assumed to arise from three latent factors: additive genetic factors (A), in which genetic effects combine additively; common or shared environment factors (C), which include all environmental exposures that are shared by the twins and contribute to their similarity; and individual-specific or unique environment factors (E), which include all environmental factors not shared by the twins plus measurement error. Because MZ twins share all their genes and DZ twins share on average 50% of their segregating genes, A contributes twice as much to the resemblance in MZ compared to DZ twins for a particular trait or disorder. By definition, MZ and DZ twins share all their C factors and none of their E factors. Non-additive genetic factors such as dominance or epistasis may be parameterized as an alternative to C, with which they are confounded. This model would fit more poorly in this study because the DZ correlations are not less than half of those of the MZ.

Model fitting was performed using the software package Mx (Neale et al. 1999), the most commonly used program for twin analyses. To test the degree to which the covariation between the three cluster C PDs resulted from common factors, we applied a single-factor independent pathway model containing three common latent variables (A_C, C_C and E_C) in addition to three disorder-specific variables (A_S, C_S and E_S) for each PD. The degree to which the cluster C PDs share genetic and environmental risk factors will be reflected in the loadings on the common versus disorder-specific factors. We chose a single common factor model for two reasons. First, because it instantiates the DSM construct of cluster C, that is the degree to which all three cluster C PDs share common genetic and environmental risk factors versus disorder-specific factors. Second, for statistical reasons, with only three disorders, models with two common factors are not identified.

A full model, including all latent variables and with different parameters specified for males and females, was tested against nested sub-models with reduced numbers of parameters. The fit of the alternative models can be compared using the difference in twice the log likelihood (2 ln L), which, under certain regularity conditions, is asymptotically distributed as χ^2 with degrees of freedom (df) equal to the difference in the number of parameters (Δχ^2 test). According to the principle of parsimony, models with fewer parameters are preferable if they do not result in a significant deterioration of fit. A useful index of parsimony is the Akaike Information Criterion (AIC), which is calculated as Δχ^2 − 2Δdf (Akaike, 1987). A lower AIC value indicates superior fit.

A basic assumption in traditional twin analyses is that MZ and DZ twins are equally correlated in their exposure to trait-relevant environments. We tested the validity of this ‘equal environment assumption’ by applying polychotomous logistic regression controlling for the correlational structure of our data using independent estimating equations as operationalized in the SAS procedure GENMOD (SAS Institute, 2005). Two variables that reflected, respectively, similarity of childhood (number of years that the twins were in the same class at school and the years the twins lived in the same residence) and adult environments (frequency of in-person and telephone contact during the past year and the distance between their current residences) were constructed. In same-sex pairs, we tested whether the PD score in twin 1 interacted with our measure of environmental similarity in predicting the relevant PD score in twin 2 (dependent variable). We controlled for main effects of zygosity, sex, age and level of environmental similarity as well as shared environment effects and genetic effects. None of the six analyses testing the impact of environmental similarity on twin resemblance for AVPD, DEPD and OCPD approached significance (all p values > 0.10).

RESULTS

Prevalence and co-occurrence
Prevalence rates for categorical DSM-IV cluster C PD diagnoses were 2.1% (n = 59) for AVPD (males 1.4%, females 2.5%), 0.3% (n = 7) for
DEPD (males 0.2%, females 0.3%) and 2.5% (n = 69) for OCPD (males 2.5%, females 2.4%). The mean (S.D.) numbers of criteria (≥1) met for the cluster C PDs were: AVPD, 0.95 (1.40); DEPD, 0.76 (1.17); OCPD, 1.93 (1.62). The proportion of individuals who endorsed none, 1 or 2 or more criteria were: AVPD, 55.0, 20.6 and 24.4%; DEPD, 57.7, 23.3, and 9.0%; OCPD, 22.7, 23.4 and 53.9% respectively. Females endorsed a significantly higher number of criteria for DEPD ($\chi^2_{20} = 20, 25$, df = 8, $p = 0.009$) and OCPD ($\chi^2_{22} = 22, 61$, df = 8, $p = 0.004$) but not for AVPD ($\chi^2 = 8, 79$, df = 7, $p = 0.27$).

The phenotypic (within-individual) correlations based on dimensional representations of the PDs (Table 1), indicate substantial co-occurrence of AVPD and DEPD and significantly lower correlation between OCPD and the other two PDs for both males and females.

### Model fitting

The results of model fitting are shown in Table 2. Based on results from the univariate analyses, we used an ACE model with only quantitative sex-effects as the multivariate model against which nested submodels were compared (model I). Specifying equal parameters for males and females resulted in a non-significant deterioration in fit and an increase in AIC (model II), and the subsequent models were therefore fitted without sex-specific effects. An AE model (without common or specific C) fitted the data well (model III), whereas a CE model without any genetic effects (model IV) was fitted by the $\chi^2_{15}$ test (40.78, $p < 0.001$), indicating a significant contribution by additive genetic effects on cluster C PDs. Models without common or specific Cs (model V and VI) were both compatible with the data, whereas a model with no specific A (model VII) was rejected by the $\chi^2_{12}$ test (29.14, $p = 0.004$), indicating that the genetic effects on cluster C PDs are partly specific to each disorder.

The parameter estimates for the best-fitting model (AE, model III) are shown in Fig. 1. Genetic effects accounted for 35% of the variance in AVPD, 31% of the variance in DEPD and 27% of the variance in OCPD. The common genetic factor accounted for 83% of the genetic influence on AVPD, 48% in DEPD and 15% in OCPD. Figure 2 summarizes the proportion of variance accounted for by common and PD-specific genetic and environmental factors. Common A and E factors accounted for 54% of the variance in AVPD and 64% of the variance in DEPD but only 11% of the variance in OCPD, indicating that OCPD is mostly etiologically distinct from the two other cluster C PDs.

### DISCUSSION

To our knowledge, this is the first population-based study of the genetic and environmental influences on DSM-IV cluster C PDs and their inter-relationship.

#### Genetic and environmental risk factors in males and females

Familial aggregation of a trait or a disorder can be caused by genetic and/or shared environmental factors. Our results indicate that for AVPD, DEPD and OCPD, familiality is best explained by genetic factors alone, with moderate genetic influence on all three PDs. Given the moderate size of our sample and thus our limited power (Neale et al. 1994; Sullivan & Eaves, 2002) we cannot rule out shared environmental effects. However, in the only twin study of DSM PDs previously published, none of the best-fitting models included shared environment (Torgersen et al. 2000). Heritability estimates for DSM-III-R AVPD, DEPD and OCPD in that study, including mostly patients with severe psychiatric disorders, were 0.28, 0.57 and 0.77 respectively. Confidence intervals were not presented, but the small sample size suggests that they would have been wide. Our results can also usefully be compared to estimates from other

---

Table 1. Phenotypic correlations based on dimensional representations of DSM-IV Cluster C personality disorders (PDs) in males and females

<table>
<thead>
<tr>
<th></th>
<th>Avoidant PD</th>
<th>Dependent PD</th>
<th>Obsessive-compulsive PD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avoidant PD</td>
<td>*</td>
<td>0.50 (0.44–0.56)</td>
<td>0.22 (0.15–0.29)</td>
</tr>
<tr>
<td>Dependent PD</td>
<td>0.59 (0.55–0.63)</td>
<td>*</td>
<td>0.25 (0.18–0.32)</td>
</tr>
<tr>
<td>Obsessive-compulsive PD</td>
<td>0.25 (0.20–0.31)</td>
<td>0.26 (0.21–0.31)</td>
<td>*</td>
</tr>
</tbody>
</table>

a Results for men are depicted above, and for women below, the diagonal formed by the asterisks.
conceptualizations of PDs. Livesley et al. (1993) and Jang et al. (1996) studied dimensions and facets of PD traits in samples of volunteer twin pairs from the general population. In the largest sample (Jang et al. 1996), no evidence was found for shared environmental effects. Heritability estimates ranged from 0.25 to 0.53 for traits related to cluster C PDs. Numerous studies have shown that DSM PDs can be represented by other models (Trull, 2005), for example the Five-Factor Model (FFM) of normal personality (Widiger & Costa, 2002). The domains and facets related to cluster C PDs in the most popular operationalization of the FFM, the NEO-Personality Inventory Revised (NEO-PI-R; Costa & McCrae, 1992), have been shown to be heritable in the range 29–46%, with no evidence of shared environmental effects (Jang et al. 1998). Our heritability estimates thus appear to be in the low end of those previously reported for personality traits resembling cluster C PDs. Our PD measures include a smaller number of items than most personality trait assessments, and may therefore include a greater proportion of measurement error (Neale et al. 2005), which may explain this result.

No evidence was found for any sex differences in genetic and environmental influences on

Table 2. Multivariate model-fitting results

<table>
<thead>
<tr>
<th>Model</th>
<th>Sex effects</th>
<th>AC</th>
<th>Cc</th>
<th>Ec</th>
<th>AS</th>
<th>Es</th>
<th>χ²</th>
<th>df</th>
<th>p</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>++</td>
<td>0.54</td>
<td>0.50</td>
<td>0.39</td>
<td>0.70</td>
<td>0.21</td>
<td>0.26</td>
<td>16.67</td>
<td>9</td>
<td>0.05</td>
</tr>
<tr>
<td>II</td>
<td>+</td>
<td>0.40</td>
<td>0.45</td>
<td>0.48</td>
<td>0.81</td>
<td>0.24</td>
<td>0.64</td>
<td>17.41</td>
<td>15</td>
<td>0.30</td>
</tr>
<tr>
<td>III a</td>
<td>+</td>
<td>0.39</td>
<td>0.29</td>
<td>0.56</td>
<td>0.80</td>
<td>0.29</td>
<td>0.47</td>
<td>40.78</td>
<td>15</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>IV</td>
<td>−</td>
<td>0.39</td>
<td>0.29</td>
<td>0.56</td>
<td>0.80</td>
<td>0.29</td>
<td>0.47</td>
<td>17.88</td>
<td>12</td>
<td>0.12</td>
</tr>
<tr>
<td>V</td>
<td>−</td>
<td>0.24</td>
<td>0.45</td>
<td>0.48</td>
<td>0.81</td>
<td>0.24</td>
<td>0.64</td>
<td>16.81</td>
<td>12</td>
<td>0.16</td>
</tr>
<tr>
<td>VI</td>
<td>+</td>
<td>0.39</td>
<td>0.29</td>
<td>0.56</td>
<td>0.80</td>
<td>0.29</td>
<td>0.47</td>
<td>29.14</td>
<td>12</td>
<td>0.004</td>
</tr>
<tr>
<td>VII</td>
<td>+</td>
<td>0.54</td>
<td>0.50</td>
<td>0.39</td>
<td>0.70</td>
<td>0.21</td>
<td>0.26</td>
<td>0.29</td>
<td>0.15</td>
<td>0.07</td>
</tr>
</tbody>
</table>

a Best-fitting model.

AC and AC, additive genetic effects; Cc and Cc, shared environmental effects; Ec and Ec, unique environmental effects; df, degrees of freedom; AIC, Akaike’s Information Criterion; +, factor estimated in model; −, factor set to zero or constrained in the model.

Fig. 1. Best-fitting model with parameter estimates and 95% confidence intervals. AC and EC, common additive genetic and individual environmental factors; AS and ES, disorder-specific genetic and individual environmental factors.

Fig. 2. Proportion of variance accounted for by common and specific genetic and environmental factors. AC (■) and EC (●), common additive genetic and individual environmental factors; AS (□) and ES (□), disorder-specific genetic and individual environmental factors; PD, personality disorder. Proportion of variance accounted for by common genetic and environmental factors is below the dotted line.
cluster C PDs. Sex differences were not explored in any of the above-mentioned studies. Evidence has, however, been found for quantitative sex differences in the heritability of Neuroticism (Lake et al., 2000), a trait closely related to AVPD and DEPD (Dyce & O’Connor, 1998). For Axis I anxiety disorders, sex differences in genetic effects have not been reported (Hettema et al., 2005). With our sample size and level of measurement we do not have the statistical power to conclude with confidence that sex effects were not present.

Common genetic and environmental risk factors

A common genetic factor accounted for most of the genetic variance in AVPD and about half of the genetic variance in DEPD, consistent with findings from the only family study of cluster C PDs that showed a close familial relationship between AVPD and DEPD (Reich, 1989).

Genetic influences on OCPD were mostly specific to this PD. This is broadly consistent with results from studies using alternative conceptualizations of PDs. The multivariate genetic analyses of lower order PD traits by Livesley et al. (1998) yielded four genetic factors that were remarkably similar to the phenotypic factors identified by principal component analysis. The lower-order trait Compulsivity, which resembles OCPD, appeared to be distinct from other traits both genetically and phenotypically. Anxiousness and social avoidance associated with AVPD and submissiveness and insecure attachment associated with DEPD were phenotypically and genetically related to the same higher-order factor, Emotional Dysregulation. In a population-based sample, Dyce & O’Connor (1998) found that DSM AVPD and DEPD correlated strongly positively with to NEO-PI-R Neuroticism and weakly negatively with Conscientiousness. By contrast, OCPD correlated strongly positively with Conscientiousness and weakly negatively with Neuroticism (Dyce & O’Connor, 1998). McCrae et al. (2001) have shown that the five-factor structure of the NEO-PI-R was found on both the phenotypic and genetic level, indicating that genetic influences on both Neuroticism and Conscientiousness are highly trait specific.

In addition to indicating that our constructs are not identifying unique genetic liabilities, the results suggest that some of the same individual environmental experiences influence different PD traits, that is environmental effects are not specific. Common unique environmental factors accounted for most of the environmental variance in DEPD and more than one-third of the environmental variance in AVPD. However, similar to the pattern found for genetic effects, unique environmental factors influencing OCPD were mostly specific to this PD. From this type of study it is not possible to determine which genetic or environmental factors may be involved.

Our results do not provide support for the validity of the DSM-IV cluster C construct in its present form. Given that several phenotypic studies indicate that OCPD stands apart from the other DSM clusters (Kass et al. 1985; Hyler & Lyons, 1988; Sanislow et al. 2002) and twin studies of personality traits show that the phenotypic structure closely reflects the underlying genetic structure (Livesley et al. 1998; Krueger, 2000; McCrae et al. 2001; Livesley, 2005), our results can be viewed as supporting the hypothesis that OCPD represents a separate Axis II secondary domain.

Limitations

The results from our study should be interpreted in light of several limitations. First, because of the low prevalence, we were unable to analyze the categorical PD diagnoses. To increase power, we instead examined dimensional representations of the DSM-IV diagnoses conceptualized as the number of criteria (≥1) endorsed. As twin analysis are based on the liability threshold model (Falconer, 1965), it should in principle make no difference if the variable studied is categorical or dimensional as long as the dimensional variable reflects the same underlying liability as the categorical diagnosis. We supported this assumption using multiple threshold tests for each individual criterion and for the dimensional representations of the three PDs. We also compared our model fitting results for OCPD in females (where the prevalence of criteria ≥2 was sufficiently high) and found that the parameter estimates were almost identical. Dimensional representations of DSM-IV PDs (Oldham & Skodol, 2000) have been shown to predict functional impairment as well as categorical diagnoses (Skodol et al. 2005).
Second, although we included a large number of twins, substantial attrition was observed in this sample from the birth registry through three waves of contact consisting of two questionnaires and a personal interview. We will report detailed analyses of the predictors of non-response across waves elsewhere (Harris et al. unpublished observations). In brief, cooperation was strongly and consistently predicted by female sex, monozygosity, older age and higher educational status, but not symptoms of mental disorder. In particular, we assessed PD traits at the second questionnaire with 91 self-report items. We used these items to predict the PD scores from the interview. The polychoric correlations between the scores based on the questionnaires and those from the interview were 0.60 for AVPD, 0.49 for DEPD, and 0.35 for OCPD. Controlling for demographic variables, these weighted scores from the second-wave questionnaire did not predict participation in the personal interview (all p > 0.20). While we cannot be certain that our sample was representative with respect to cluster C psychopathology, these findings suggest that a substantial bias is unlikely. Third, the twins were interviewed only once. Although we demonstrated high inter-rater reliability and internal consistency, we could not estimate the test–retest reliability over time. Previous studies have shown that the 2-year test–retest reliability of AVPD and OCPD is relatively low (McGlashan et al. 2005). In twin analyses, measurement errors are reflected in E, which implies that a reduction in reliability would result in decreased heritability estimates. Furthermore, analyses of sum scores may yield biased estimates of variance components of the latent trait. Thus, the analyses reported here may be subject to these biases, which are likely to deflate the familial (A and C) and inflate the non-familial (E) components (Neale et al. 2005). Finally, these results were obtained from a sample of young Norwegian adults, and may or may not extrapolate to other age cohorts or ethnic groups. However, prevalence estimates from a recent Norwegian epidemiological study in a community sample were within the same range as those reported from community studies in other western countries (Torgersen et al. 2001). The participants in our sample were twins. A previous study of personality failed to show any systematic differences between twin and non-twin samples (Johnson et al. 2002).

ACKNOWLEDGMENTS

This work was supported by grant MH-068643 from the National Institutes of Health. The twin program of research at the Norwegian Institute of Public Health is supported by grants from the Norwegian Research Council and the Norwegian Foundation for Health and Rehabilitation, and by the European Commission under the program ‘Quality of Life and Management of the Living Resources’ of the 5th Framework Program (no. QLG2-CT-2002-01254). Genotyping on the twins was performed at the Starr Genotyping Resource Centre at Rockefeller University. We are very grateful to the twins for their participation.

DECLARATION OF INTEREST

None.

REFERENCES


The structure of genetic and environmental risk factors for phobias in women

Nikolai Czajkowski MA, MS
Kenneth S. Kendler MD
Kristian Tambs PhD
Espen Røysamb PhD
Ted Reichborn-Kjennerud MD

From the Division of Mental Health, Norwegian Institute of Public Health (N. Czajkowski, K. Tambs, E. Røysamb, T. Reichborn-Kjennerud), The Virginia Institute for Psychiatric and Behavioral Genetics and Department of Psychiatry and Human Genetics Medical College of Virginia/Virginia Commonwealth University, Richmond, VA. USA (K. Kendler)
Department of Psychology, University of Oslo (E. Røysamb)

Corresponding author: Nikolai Czajkowski, Division of Mental Health, Norwegian Institute of Public Health, Box 4404, Nydalen N- 0403 Oslo, Norway.
Phone: (+47) 21 07 82 48
Fax :  (+47) 21 07 81 01
e-mail: Nikolai.Czajkowski@fhi.no
Abstract

**Background:** To explore the genetic and environmental factors underlying the co-occurrence of lifetime diagnoses of DSM-IV phobia.

**Methods:** Female twins (n=1430) from the population-based Norwegian Institute of Public Health Twin Panel were assessed at personal interview for DSM-IV lifetime specific phobia, social phobia and agoraphobia. Comorbidity between the phobias were assessed by odds-ratios and polychoric correlations, and multivariate twin models were fitted in Mx.

**Results:** Phenotypic correlations of lifetime phobia diagnoses ranged from 0.55 (agoraphobia and social phobia, OR=10.95) to 0.06 (animal phobia and social phobia, OR=1.21). In the best fitting twin model, which did not include shared environmental factors, heritability estimates for the phobias ranged from 0.43 to 0.63. Comorbidity between the phobias was accounted for by two common liability factors. The first loaded principally on animal phobia, but also weakly on the other specific phobias, but did not influence the complex phobias (agoraphobia and social phobia). The second liability factor strongly influenced the complex phobias, but also loaded weakly to moderately on all the other phobias. Blood phobia was mainly influenced by a specific genetic factor which accounted for 51% of the total and 81% of the genetic variance.

**Conclusions:** Phobias are highly comorbid and heritable. Our results suggest that the comorbidity between phobias is best explained by two distinct liability factors rather than a single factor, as has been assumed in most previous multivariate twin analyses. One of these factors was specific to the simple phobias, while the other was more general. Blood phobia was mainly influenced by disorder specific genetic factors.
A population based family study of symptoms of anxiety and depression. A HUNT study.

Nikolai Czajkowski M.A., M.S.
Espen Røysamb PhD
Ted Reichborn-Kjennerud MD, PhD
Kristian Tambs PhD

From the Division of Mental Health, Norwegian Institute of Public Health (N. Czajkowski, E. Røysamb, T. Reichborn-Kjennerud, K. Tambs),
Department of Psychology, University of Oslo (E. Røysamb)

Corresponding author: Nikolai Czajkowski, Division of Mental Health, Norwegian Institute of Public Health, Box 4404, Nydalen N- 0403 Oslo, Norway.
Phone: (+47) 21 07 82 48
Fax : (+47) 21 07 81 01
e-mail: Nikolai.Czajkowski@fhi.no
Objective: To estimate an upper limit on the heritability of self reported symptoms of anxiety and depression in a large and population representative nuclear family sample.

Method: The ten-item symptom checklist (SCL-10) was administered as part of a health survey in a Norwegian county. The SCL-10 is a shortened version of the SCL-25, assessing symptoms of anxiety and depression. In all, 46,064 people responded, and with data from Statistics Norway, responses of first degree relatives could be linked. Polychoric correlations between family members score on SCL-10 were calculated, and a structural equation model was fitted to these correlations using the software package R.

Results: All correlations between nuclear family members were in the range 0.12 to 0.16, indicating small but significant familial influences on SCL-10. In the best fitting model, heritability was estimated at 0.25 (95% CI = 0.22-0.27), and sibling specific environmental effects could be discarded.

Conclusions: The estimated upper level heritability for SCL-10 in our sample was lower than what has been reported in twin studies of similar measures.
Appendix I

DSM-IV-TR diagnostic criteria for PAPD and Cluster C PDs
DSM-IV-TR Criteria for Passive-Aggressive personality disorder

A. A pervasive pattern of negativistic attitudes and passive resistance to demands for adequate performance, beginning by early adulthood and present in a variety of contexts, as indicated by four (or more) of the following:

(1) passively resists fulfilling routine social and occupational tasks;
(2) complains of being misunderstood and unappreciated by others;
(3) is sullen and argumentative;
(4) unreasonably criticizes and scorns authority;
(5) expresses envy and resentment toward those apparently more fortunate;
(6) voices exaggerated and persistent complaints of personal misfortune;
(7) alternates between hostile defiance and contrition.

B. The disorder does not occur exclusively during Major Depressive Episodes and is not better accounted for by Dysthymic Disorder.
DSM-IV-TR Criteria for Avoidant personality disorder

A Pervasive pattern of social inhibition, feelings of inadequacy, and hypersensitivity to negative evaluation, beginning by early adulthood and present in a variety of contexts, as indicated by four (or more) of the following:

(1) Avoids occupational activities that involve significant interpersonal contact, because of fears of criticism, disapproval, or rejection

(2) Is unwilling to get involved with people unless certain of being liked

(3) Shows restraint initiating intimate relationships because of the fear of being ashamed, ridiculed, or rejected due to severe low self-worth.

(4) Is preoccupied with being criticized or rejected in social situations

(5) Is inhibited in new interpersonal situations because of feelings of inadequacy

(6) Views self as socially inept, personally unappealing, or inferior to others

(7) Is unusually reluctant to take personal risks or to engage in any new activities because they may prove embarrassing
DSM-IV-TR Criteria for Obsessive Compulsive personality disorder

A. A pervasive pattern of preoccupation with orderliness, perfectionism, and mental and interpersonal control, at the expense of flexibility, openness, and efficiency, beginning by early adulthood and present in a variety of contexts, as indicated by four (or more) of the following:

(1) Is preoccupied with details, rules, lists, order, organization, or schedules to the extent that the major point of the activity is lost.

(2) Shows perfectionism that interferes with task completion (e.g., is unable to complete a project because his or her own overly strict standards are not met)

(3) Is excessively devoted to work and productivity to the exclusion of leisure activities and friendships (not accounted for by obvious economic necessity)

(4) Is overconscientious, scrupulous, and inflexible about matters of morality, ethics, or values (not accounted for by cultural or religious identification)

(5) Is unable to discard worn-out or worthless objects even when they have no sentimental value.

(6) Is reluctant to delegate tasks or to work with others unless they submit to exactly his or her way of doing things

(7) Adopts a miserly spending style toward both self and others; money is viewed as something to be hoarded for future catastrophes

(8) Shows rigidity and stubbornness
DSM-IV-TR Criteria for Dependent Personality Disorder

A. A pervasive and excessive need to be taken care of that leads to submissive and clinging behavior and fears of separation, beginning by early adulthood and present in a variety of contexts, as indicated by five (or more) of the following:

(1) has difficulty making everyday decisions without an excessive amount of advice and reassurance from others

(2) needs others to assume responsibility for most major areas of his or her life

(3) has difficulty expressing disagreement with others because of fear of loss of support or approval.

(4) has difficulty initiating projects or doing things on his or her own (because of a lack of self-confidence in judgment or abilities rather than a lack of motivation or energy)

(5) goes to excessive lengths to obtain nurturance and support from others, to the point of volunteering to do things that are unpleasant

(6) feels uncomfortable or helpless when alone because of exaggerated fears of being unable to care for himself or herself

(7) urgently seeks another relationship as a source of care and support when a close relationship ends

(8) is unrealistically preoccupied with fears of being left to take care of himself or herself
Appendix II

DSM-IV-TR diagnostic criteria for phobias
DSM-IV-TR criteria for specific phobia

A. Marked and persistent fear that is excessive or unreasonable, cued by the presence or anticipation of a specific object or situation (e.g., flying, heights, animals, receiving an injection, seeing blood).

B. Exposure to the phobic stimulus almost invariably provokes an immediate anxiety response, which may take the form of a situationally bound or situationally predisposed Panic Attack. Note: In children, the anxiety may be expressed by crying, tantrums, freezing, or clinging.

C. The person recognizes that the fear is excessive or unreasonable. Note: In children, this feature may be absent.

D. The phobic situation(s) is avoided or else is endured with intense anxiety or distress.

E. The avoidance, anxious anticipation, or distress in the feared situation(s) interferes significantly with the person's normal routine, occupational (or academic) functioning, or social activities or relationships, or there is marked distress about having the phobia.

F. In individuals under age 18 years, the duration is at least 6 months.

G. The anxiety, Panic Attacks, or phobic avoidance associated with the specific object or situation are not better accounted for by another mental disorder, such as Obsessive-Compulsive Disorder (e.g., fear of dirt in someone with an obsession about contamination), Posttraumatic Stress Disorder (e.g., avoidance of stimuli associated with a severe stressor), Separation Anxiety Disorder (e.g., avoidance of school), Social Phobia (e.g., avoidance of social situations because of fear of embarrassment), Panic Disorder With Agoraphobia, or Agoraphobia Without History of Panic Disorder.
DSM-IV-TR criteria for social phobia

According to the DSM-IV-TR, to be diagnosed with Social Phobia all these criteria (A-H) must be met:

A. A marked and persistent fear of one or more social performance situations in which the person is exposed to unfamiliar people or to possible scrutiny by others. The individual fears that he or she will act in a way (or show anxiety symptoms) that will be humiliating or embarrassing. Note: In children, there must be evidence of the capacity for age-appropriate social relationships with familiar people and the anxiety must occur in peer settings, not just in interactions with adults.

B. Exposure to the social or performance situation almost invariably provokes an immediate anxiety response. This response may take the form of a situationally bound or situationally people predisposed Panic Attack. Note: In children, the anxiety may be expressed by crying, tantrums, freezing, or shrinking from social situations with unfamiliar people.

C. The person recognizes that their fear is excessive or unreasonable. Note: In children, this feature may be absent.

D. The social or performance situation is avoided, although it is sometimes endured with dread (intense anxiety or distress).

E. The avoidance, anxious anticipation of, or distress in, the feared social or performance situation interferes significantly with the person's normal routine, occupational (academic) functioning, social life, or if the person is markedly distressed about having the phobia.

F. In individuals under age 18 years, the duration is at least 6 months.

G. The fear or avoidance is not due to the direct physiological effects of a substance or a general medical condition and is not better accounted for by another mental disorder (e.g., Panic Disorder, Separation Anxiety Disorder, Body Dysmorphic Disorder, a Pervasive Developmental Disorder, or Schizoid Personality Disorder).

H. If a general medical condition or another mental disorder is present, the fear in Criterion A or the avoidance in Criteria D, is unrelated to it (e.g., the fear is not of Stuttering, trembling in Parkinson's disease, or exhibiting abnormal eating behavior in Anorexia Nervosa).

Specify if:
Generalized: if the fears include most social situations (also consider the additional diagnosis of Avoidant Personality Disorder).
DSM-IV-TR diagnostic criteria for agoraphobia

A. Anxiety about being in places or situations from which escape might be difficult (or embarrassing) or in which help may not be available in the event of having an unexpected or situationally predisposed panic attack or panic-like symptoms. Agoraphobic fears typically involve characteristic clusters of situations that include being outside the home alone; being in a crowd, or standing in a line; being on a bridge; and traveling in a bus, train, or automobile.

B. The situations are avoided (e.g., travel is restricted) or else are endured with marked distress or with anxiety about having a panic attack or panic-like symptoms, or require the presence of a companion.

C. The anxiety or phobic avoidance is not better accounted for by another mental disorder, such as Social Phobia (e.g., avoidance limited to social situations because of fear of embarrassment), Specific Phobia (e.g., avoidance limited to a single situation like elevators), Obsessive-Compulsive Disorder (e.g., avoidance of dirt in someone with an obsession about contamination), Posttraumatic Stress Disorder (e.g., avoidance of stimuli associated with a severe stressor), or Separation Anxiety Disorder (e.g., avoidance of leaving home or relatives).
Appendix III

SCL-10 items
<table>
<thead>
<tr>
<th>Sudden symptoms</th>
<th>Not at all</th>
<th>A little</th>
<th>Quite a bit</th>
<th>Extremely</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suddenly scared for no reason</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Nervousness or shakiness inside</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Faintness, dizziness, or weakness</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Feeling tense or keyed up</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Blaming yourself for things</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Difficulty falling asleep, staying asleep</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Feeling blue</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Feeling of worthlessness</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Feeling everything is an effort</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Feeling hopeless about the future</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

**Anxiety**  
**Depression**
Appendix IV

Example R script from family analysis of SCL-10
N_HUNT <- matrix (scan
("H:\\PhD\\HUNT\\NIK_SREL_data_SCL10_31082008_ALDER95_MASTER2_4var.dat", skip=1, na.strings="-
1"), ncol=7, byrow=T)

# ************************************************
# Threshold items to use polychoric correlations.
# ************************************************
require(polycor)

# -----------------------
# Spouses
Tr_F_M_P_P  <- polychor(MF[,1]==99,6, MF[,1]==99,7, std.err = TRUE)

# -----------------------
# Child Mother
Tr_M_CM_P_P  <- polychor(FM[,1]==1,6, FM[,1]==1,7, std.err = TRUE)
Tr_M_CF_P_P  <- polychor(FF[,1]==1,6, FF[,1]==1,7, std.err = TRUE)

# -----------------------
# Child Father
Tr_F_CM_P_P  <- polychor(MM[,1]==2,6, MM[,1]==2,7, std.err = TRUE)
Tr_F_CF_P_P  <- polychor(MF[,1]==2,6, MF[,1]==2,7, std.err = TRUE)

# -----------------------
# Correlate Children
Tr_C1M_C2M_P_P <- polychor(MM[,1]==7,6, MM[,1]==7,7, std.err = TRUE)
Tr_C1M_C2F_P_P <- polychor(MF[,1]==7,6, MF[,1]==7,7, std.err = TRUE)
Tr_C1F_C2F_P_P <- polychor(FF[,1]==7,6, FF[,1]==7,7, std.err = TRUE)

# ***************************************************************************
# Få ut estimatene
# ***************************************************************************

# Spouses
r_F_M_P_P  <- Tr_F_M_P_P$rho # Subtract Age effect

# Child Mother
r_M_CM_P_P  <- Tr_M_CM_P_P$rho  # Subtract Age effect
r_M_CF_P_P  <- Tr_M_CF_P_P$rho  # Subtract Age effect

# Child Father
r_F_CM_P_P  <- Tr_F_CM_P_P$rho# Subtract Age effect
r_F_CF_P_P  <- Tr_F_CF_P_P$rho # Subtract Age effect

# Correlate Children
r_C1M_C2M_P_P <- Tr_C1M_C2M_P_P$rho  # Subtract Age effect
# Subtract Age effect
r_C1M_C2F_P_P <- Tr_C1M_C2F_P_P$rho
r_C1F_C2F_P_P <- Tr_C1F_C2F_P_P$rho

# **************************************************************************
# Get weights
# -----------------------
# Spouses
v_F_M_P_P  <-  Tr_F_M_P_P$var
# -----------------------
# Child Mother
v_M_CM_P_P  <-  Tr_M_CM_P_P$var
v_M_CF_P_P  <-  Tr_M_CF_P_P$var
# -----------------------
# Child Father
v_F_CM_P_P  <-  Tr_F_CM_P_P$var
v_F_CF_P_P  <-  Tr_F_CF_P_P$var
# -----------------------
# Correlate Children
v_C1M_C2M_P_P <- Tr_C1M_C2M_P_P$var
v_C1M_C2F_P_P <- Tr_C1M_C2F_P_P$var
v_C1F_C2F_P_P <- Tr_C1F_C2F_P_P$var
# **********************************************************
# The main function that is to be minimized
# ********************************************************************************

my_DWLS.helper.SexSpecific.full <- function (params)
{
  DWLS<-0;
  hf<-params[1] # sqrt(HERITABILITY) females
  hm<-params[2] # sqrt(HERITABILITY) males
  Sf<-params[3] # Sibling effect
  Sm<-params[4] # Sibling effect
  Z<-params[5] # OS sibling sex effects
  M<-params[6] # E correlation between parents
  e_squared_m<-(1-hm^2);  # Initially,E and H only. E is a function of H
  e_squared_f<-(1-hf^2);  # Initially,E and H only. E is a function of H
  # **********************************************************
  # Calculate Expected correlations
  t_r_F_M_P_P  <- M;
  # ------------
  # Mother child
t_r_M_CM_P_P  <- 0.5*hm*hf)+(0.5*hm*hm*M);
  t_r_M_CF_P_P  <- 0.5*hf+hf)+(0.5*hm*hf*M);
  # ------------
  # Father child
t_r_F_CM_P_P  <- 0.5*hm*hm)+(0.5*hm*hf*M);
  t_r_F_CF_P_P  <- 0.5*hf+hf)+(0.5*hf*M);
  # ------------
  # Child child
t_r_C1M_C2M_P_P <-(hm*hm*0.5^2+hm*hm*0.5^2)+(Sm^2*e_squared_m)+(2*hm*hf*hm*hm*0.5^2*M)
t_r_C1F_C2F_P_P <-(hf*hf*0.5^2+hf*hf*0.5^2)+(Sf^2*e_squared_f)+(2*hf*hf*hm*hf*0.5^2*M)
t_r_C1M_C2F_P_P <-(hm*hf*0.5^2+hm*hf*0.5^2)+(Sm*Z*sqrt(e_squared_m)*
                        sqrt(e_squared_f))+(2*hm*hm*hf*hf*0.5^2*M)

  # Calculate deviation from expected correlations
DWLS<- 0
#
# -----------------------
# Spouses

DWLS<- DWLS + (1/v_F_M_P_P) *((atanh(r_F_M_P_P) - atanh(t_r_F_M_P_P))^2)

# -----------------------
# Child Mother

DWLS<- DWLS + (1/v_M_CM_P_P) *((atanh(r_M_CM_P_P) - atanh(t_r_M_CM_P_P))^2)
DWLS<- DWLS + (1/v_M_CF_P_P) *((atanh(r_M_CF_P_P) - atanh(t_r_M_CF_P_P))^2)

# -----------------------
# Child Father

DWLS<- DWLS + (1/v_F_CM_P_P) *((atanh(r_F_CM_P_P) - atanh(t_r_F_CM_P_P))^2)
DWLS<- DWLS + (1/v_F_CF_P_P) *((atanh(r_F_CF_P_P) - atanh(t_r_F_CF_P_P))^2)

# -----------------------
# Correlate Children

DWLS<- DWLS + (1/v_C1M_C2M_P_P) *((atanh(r_C1M_C2M_P_P) - atanh(t_r_C1M_C2M_P_P))^2)
DWLS<- DWLS + (1/v_C1M_C2F_P_P) *((atanh(r_C1M_C2F_P_P) - atanh(t_r_C1M_C2F_P_P))^2)
DWLS<- DWLS + (1/v_C1F_C2F_P_P) *((atanh(r_C1F_C2F_P_P) - atanh(t_r_C1F_C2F_P_P))^2)

#print (DWLS)
#
#
# ************************************************************
# ************************************************************

HUNT.DWLS.estimation.SexSpecific.full <- function ()
{
  tmp<- nlm(my_DWLS.helper.SexSpecific.full,c(0.3,0.3,0.2,0.3,0.3,0.04),iterlim = 200);
}

HUNT_estimate.full<-HUNT.DWLS.estimation.SexSpecific.full()