From the Department of Medical Epidemiology and Biostatistics
Karolinska Institutet, Stockholm, Sweden

VIOLENT CRIME: ADDRESSING CAUSATION WITH FAMILY-BASED METHODS

Thomas Frisell

Karolinska Institutet

Stockholm 2012

Published by Karolinska Institutet. Printed by Larserics Digital Print AB

ISBN 978-91-7457-719-8

ABSTRACT

Violent crime is an important public health problem, and incurs major costs for society. The effect of interventions has so far been modest, often attributed to a research focus on risk factors for crime, but a relative lack of understanding of the causal mechanisms behind these factors. The four studies in this thesis attempt to address different aspects of the etiology of violent crime by using family-based epidemiologic methods.

It has long been known that antisocial behavior runs in families. In Paper I, a nested case-control was used to quantify the familial clustering of violent crime using a linkage of several Swedish total population registers. We were able to provide precise estimates of the familial aggregation among 1st, 2nd, and 3rd degree relatives, and also adoptive relations and spouses. Familial risks were moderate to strong, and were modified by gender, socioeconomic status, type of violent crime, and age at first conviction. Familial clustering suggests that genes and/or family environment influence the propensity for violent offending. In Paper II we attempted to estimate the relative importance of these factors by calculating the heritability in mixed probit regression. Comparing results from twin, adoptee-parent, adoptee-sibling, and sibling designs, and attempting to adjust for non-random mating, we found that about half the variation in violent offending could be attributed to genetic factors. We also found significant gender differences in the etiology of violent crime.

In Paper III, we discussed the interpretation of sibling comparison designs. Sibling comparisons have been hailed for their ability to adjust for family-shared confounders, but have received little attention from a methodological standpoint. In line with previous research in economy, we showed that these models are subject to several caveats, and that they may in some situations increase rather than decrease bias. The implications of this were acknowledged in Paper IV, where we analysed the association of general cognitive ability and violent crime, and adjusted for shared family characteristics through sibling comparison analysis. Taking measurement error and non-shared confounding into account, the results indicated that the association was partly confounded by factors shared by siblings, but that most of the association could not be explained by such factors.

Together, Papers I and II suggested that violent crime runs in families due to both genetic and environmental factors, and Paper IV offered some support for the hypothesis that intelligence may be one of the factors explaining this familial aggregation. The caveats of sibling comparisons pointed out in Paper III should be taken into account when using co-twin control and other sibling designs to address issues of causality.

LIST OF PUBLICATIONS

I. Frisell T, Lichtenstein P, Långström N.
Violent crime runs in families: a total population study of 12.5 million individuals.
Psychological Medicine. 2011; 41: 97-105.

II. Frisell T, Pawitan Y, Långström N, Lichtenstein P.
Heritability, assortative mating and gender differences in violent crime: Results from a total population sample using twin, adoption, and sibling models.
Behavior Genetics. 2012; 42: 3-18.

III. Frisell T, Öberg S, Kuja-Halkola R, Sjölander A.
Sibling comparison designs: Bias from non-shared confounders and measurement error.
Epidemiology. 2012; In press.

IV. Frisell T, Pawitan Y, Långström N.
Is the association between general cognitive ability and violent offending caused by family-level confounders?
Submitted manuscript

CONTENTS

LIST OF ABBREVIATIONS

A	Additive genetic component; breeding value
ADHD	Attention-deficit hyperactivity disorder
ASPD	Antisocial personality disorder
BRÅ	Brottsförebyggande rådet, National Council of Crime Prevention
C	Sibling (or family) shared environment
CD	Conduct disorder
CI	Confidence Interval
DAG	Directed acyclic graph
DZ	Dizygotic; fraternal
DZOS	Dizygotic of opposite sexes
E	Residual variance; Unique environment
GEE	Generalized Estimating Equation
GLM	Generalized Linear Model
GLMM	Generalized Linear Mixed Model
GWAS	Genome-Wide Association Study
IQ	Intelligence quotient; general cognitive ability
MAOA	Monoamine oxidase type A
MGR	Multi-Generation Register
MLE	Maximum Likelihood Estimate
MZ	Monozygotic; identical
NTU	Nationella trygghetsundersökningen, National Safety Survey
OR	Odds ratio
SALT	Screening Across the Lifespan Twin study
SNP	Single Nucleotide Polymorphism
STAGE	Study of Twin Adults: Genes and Environments
STR	Swedish Twin Register
WHO	World Health Organization

1 INTRODUCTION

Whether measured in monetary terms or in physical and emotional trauma, the cost of severe antisocial and violent criminal behavior is a major concern in modern society. Indeed, interpersonal violence is recognized by the World Health Organization (WHO) as an important public health problem [1], and the third leading cause of death among adolescents and young adults in Europe [2]. The occurrence of violent offending is not evenly distributed in the population; most criminal offences are committed by adolescents and young adults, and men are nearly ten times as likely as women to be convicted of a violent crime. Notably, a few per cent of the population, usually with early onset and a multitude of individual and psychosocial risk factors, commit about half of all criminal acts [3, 4]. If future persistent offenders could be identified, and their violence could be prevented, society would benefit not only from reduced victimization, but also through increased productivity from these otherwise criminal individuals. Although it is very difficult to estimate such economic costs with any degree of precision, a US study suggested that for each high risk individual that could be identified at birth and helped to develop a normative rather than criminal career, society would save in the range of $2.6- $4.4 million [5].

Many risk factors for crime have been identified, but the effects of prevention and treatment are at best moderate [6]. It has often been proposed that the effect of interventions might be considerably improved if we had a better understanding of the causal mechanisms behind these risk factors. The field of criminology has even been accused of being "stuck at the risk factor stage" [4, 6], since few studies have been able to test competing causal theories against each other. In particular, authorities in the field have expressed the need for "genetically informative designs", i.e. designs using pedigrees or measured genes, able to separate the effect of experienced environments from an individual's inherent characteristics [6, 7].

The four studies in this thesis attempt to address different aspects of the etiology of violent crime by using, and problematizing, family-based epidemiological methods. Papers I, II and IV are based on the finding that crime and other antisocial behavior cluster in families [8, 9]. There is information in this familial aggregation, beyond the use of family history as a risk marker in prediction models [10]. The magnitude of familial risks put an upper bound on the possible strength of risk factors shared in families, such as socioeconomic factors, and parenting practices. By comparing relatives at different levels of relatedness, we may also be able to estimate the relative importance of genetic and environmental factors. Sibling comparison designs capitalize on familial clustering by using siblings as controls, hoping to control for confounding by factors shared by the siblings. Paper III is a methodological critique of these models, suggesting that the complexity in interpreting results from these comparisons might often have been

underestimated, and Paper IV is an application of the design to the association of general intelligence and violent offending.

Before summarizing and discussing the results of the studies, this thesis introduction will introduce and contextualize some of the definitions, causal concepts, and statistical methods used throughout the four papers.

1.1 VIOLENT CRIME

Violent crime may be defined as criminal acts of interpersonal violence. In accordance with the WHO's definition of interpersonal violence [1], this would include psychological as well as physical abuse. Both actual hitting and convincing threats of hitting would be considered violent acts, albeit of different severity. What constitutes a criminal act depends on the jurisdiction, but with the exception of violence sanctioned by the state, e.g. used by the police and military forces, non-sexual violent acts are almost universally illegal.

1.1.1 Violent crime in Sweden

The papers in this thesis are all based on registered convictions of violent crime in Sweden. Relying on convictions for violence avoids problems with recall or response bias. It is well known, however, that only a fraction of all violent acts result in a conviction for violent offending. The Swedish National Council on Crime Prevention (Brottsförebyggande rådet, BRÅ) is the governmental agency responsible for producing and publishing crime statistics in Sweden. Since 2006, BRÅ tries to estimate the "dark figures" for different crimes by collecting information on victimization through yearly safety surveys (Nationella trygghetsundersökningen, NTU). In 2006-2007, these surveys suggested that 21% of all illegal threats, 34% of all assaults, and 43% of all robberies were reported to the police [11]. The severity of a crime was reported as the main factor in deciding whether to report it, but the perceived severity of specific criminal acts may change over time. A trend of increasing police reports of assaults, but no increase in self-reported victimization, has been observed in several West European countries [12]. There are also exceptions; sexual offences, although generally considered as serious offences, had a report rate of only 13% in the NTU [11]. Of violent crimes reported to the police, only about 20% leads to a suspected perpetrator prosecuted in court [13]. In short, only a minority of all violent crimes leads to registered convictions, but the latter are more likely to capture severe violent crimes.

The discrepancy between self-reported and official crime makes it difficult to compare crime rates across countries. Differences in both legal practice and the tendency of victims to report crime may lead to differences in official crime statistics. Based on data from the International Crime and Victimization Study (a series of interviews performed 1989-2000 in multiple European countries), Sweden is close to average among Western European countries on assaults,

threats, thefts, and sexual crime, overrepresented in bicycle thefts, and perhaps slightly below average on robberies and burglaries [12]. Homicide rates may be more reliably estimated through registered cause of death information. Based on such data, Sweden has similar rates of homicide as Denmark and Norway, an incidence of about 1/100,000 person-years, or less than 100 homicides per year in Sweden [12]. While similar to most West European countries, this is low compared to the global average. According to FBI statistics, the homicide rate in the US was 4.8/100,000 in 2010, and 9.8/100,000 in 1991 [14].

During the past 50 years, immigration to Sweden has increased. The proportion of the Swedish population born in another country increased from 4% in 1960, to 11% in 2000, and 14% in 2008 [15]. About half immigrated from Scandinavian countries and the EU. According to the official crime statistics for 1997, immigrants were 2.5 times more likely than Swedish-born to be suspected of a crime [16]. Children of immigrants were also at an increased risk, twice as likely as other Swedish-born to be suspected of a crime. The relative risks were higher for some specific crimes, with a four times increased risk for lethal violence and robbery, and a five times increased risk for rape [16]. The reasons for the over representation of crime among immigrants is relatively unexplored in the Swedish context, but is often thought likely to be due to socioeconomic differences, stress from the migration process, and deficiencies in the Swedish integrative system [16]. In one of the few studies on the topic, it was shown that being suspected of violent crime among Swedish immigrant men was predicted by the native country's level of human development, but not by its history of war [17]. If immigration is increasing, and immigrants are overrepresented in crime, it may be expected that crime would have increased in Sweden over the past decades. Assessed through deaths due to homicide, there may have been an increase 1950-1990, but in recent decades there may actually have been a decreasing trend [18]. In contrast, the number of individuals convicted for assault increased from 1970-1990 and has been stable since [19]. However, the proportion of assaults that represent violence against women and children, which may previously have gone unreported, has increased [19]. Thus, it seems that there has been an increase in violent crime until 1990, but that the rate of violent crime may have been stable, or even decreased, since then.

1.1.2 Violent crime is an antisocial behavior

Research on violence and criminal behavior define the outcome in a multitude of ways, focusing on partly overlapping but potentially different behaviors or personality styles.

Antisocial behavior is an umbrella term for externalizing (acting-out) behavior that violates the right of others and/or conflicts with established norms. It includes criminal acts, violent or otherwise, but is often used to also capture other aggressive or oppositional behavior and adjustment problems. The term *delinquency* has been used to refer to rule-breaking behavior among adolescents

or young adults. Though some define the term as criminal acts among juveniles [20], it may often contain behaviors that, depending on jurisdiction may not be strictly illegal, such as truancy or under-age drinking.

There are several psychiatric disorders that are closely connected to antisocial or criminal behavior. Among children and adolescents *Oppositional Defiant Disorder* and *Conduct Disorder* are defined by a consistent pattern of antisocial and aggressive behavior. Among adults, *Antisocial Personality Disorder* (ASPD) is defined as "a pattern of disregard for, and violation of, the rights of others" [21]. The diagnostic criteria for ASPD currently include committing illegal acts, being aggressive, irritable or impulsive, and lacking remorse. *Psychopathy* refers to a personality style demarked by grandiose narcissism, emotional detachment and lack of empathy, and antisocial behavior [22]. Whether psychopathy should be seen as a diagnostic entity separate from ASPD has been debated, but its overrepresentation in prisons and forensic settings is beyond question.

With the possible exception of the psychiatric diagnoses, the definitions vary from study to study, and over time. For instance, in a study of juvenile delinquency published in 1936, the delinquency of one of the study participants was thought evident through his unrepentant homosexual behavior [23]. Few researchers would embrace that definition today, but perhaps similarly culturally sensitive; early sexual debut or having many sexual partners are sometimes considered externalizing or even antisocial behavior.

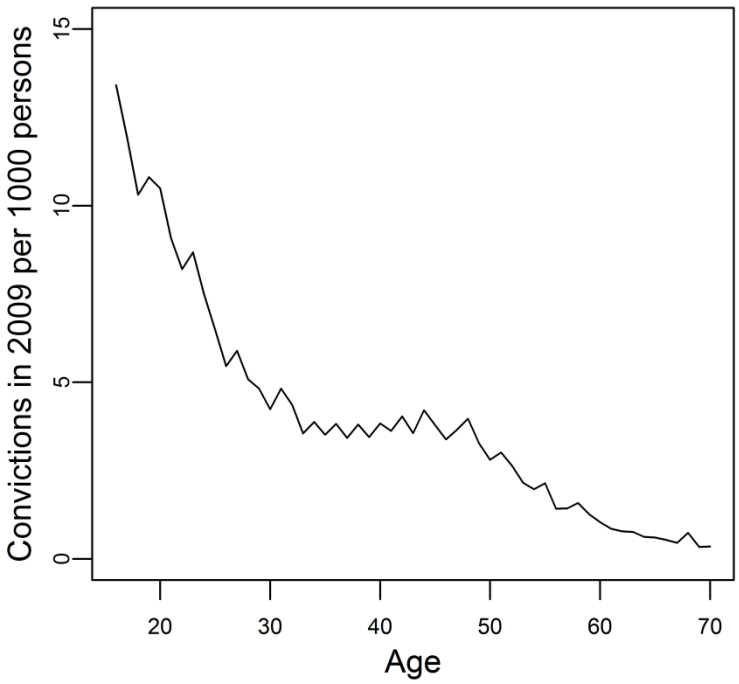
The differently defined antisocial behaviors are similar, but they do not necessarily measure the same thing. Behavior genetic studies point to differences between aggressive and non-aggressive antisocial behavior [24], reactive and proactive aggression [25, 26] and antisocial behavior at different stages of a person's life [27]. Although mental disorder is a risk factor for crime generally, it seems specifically strong for arson [28]. On the other hand, persistent offenders commit many different types of crimes [29]. Repeat offenders tend to be "versatile", committing both violent and non-violent offences [30]. It has been suggested that while some sexual offenders seem to be versatile offenders, others, in particular those targeting children, may be more specialized, and should be understood in light of paraphilic sexual preferences [31]. This is supported by differences reported in childhood risk factors among sexual versus non-sexual violent offenders [32], and adolescents reporting sexual versus non-sexual conduct problems [33]. Overall, behavior genetic modeling offers support for both considerable etiological overlap between different antisocial behavior constructs, and unique factors and dynamics influencing single traits [34-36].

1.1.3 Antisocial behavior changes over the life-course

There is evidence for both stability and change in antisocial behavior over an individual's life time [37]. Aggression as early as age 3 has been shown to significantly predict adolescent aggression [38]. Childhood conduct disorder is a

relatively strong predictor of juvenile delinquency [39], criminal offending [40], and partner violence [41]. Conduct disorder is also a prerequisite for the diagnosis of adult antisocial personality disorder [42]. Among adults, the strongest risk factors for criminal recidivism is an "antisocial personality" or history of previous crime [43]. Despite these signs of stability, the rate of antisocial behavior varies with age, peaking in adolescence or early adulthood. This observation led to Moffitt's influential developmental taxonomy of antisocial behavior, where she posited two distinct types of offenders, the adolescence-limited type and the life-course persistent type [44]. Adolescent-limited offenders would be influenced by peer influences and age-specific norms during a developmental period marked by physiological and psychological changes, and desist from antisocial behavior as they mature. Life-course persistent offenders would start their delinquent behavior in early childhood, influenced by neuropsychological problems in combination with criminogenic environments. Moffitt's taxonomy has been very influential, and there is support that individuals showing earlier antisocial behavior and/or committing more serious offences have lower intelligence [29, 45-48], come from more troubled homes [29, 45, 46, 49-51], has more neuropsychiatric problems [29, 47, 52], and antisocial relatives [29, 50]. However, it has also been shown that the peak ages vary with different antisocial behaviors. Aggression seems to peak in early childhood [38, 53], while several criminal offences peak later (cf Figure 1). When these types of antisocial behavior are taken into account, the peak in adolescence is less distinct. Attempts at modeling trajectories of antisocial behavior have found some support for Moffitt's adolescence-limited type, but often additionally find other distinct trajectories, e.g. "high" versus "low" persistent offenders [29, 49, 54, 55], or "late-bloomers", persistent or serious offenders with a late age of initiation [52, 56].

Fig1. Age distribution of violent crime

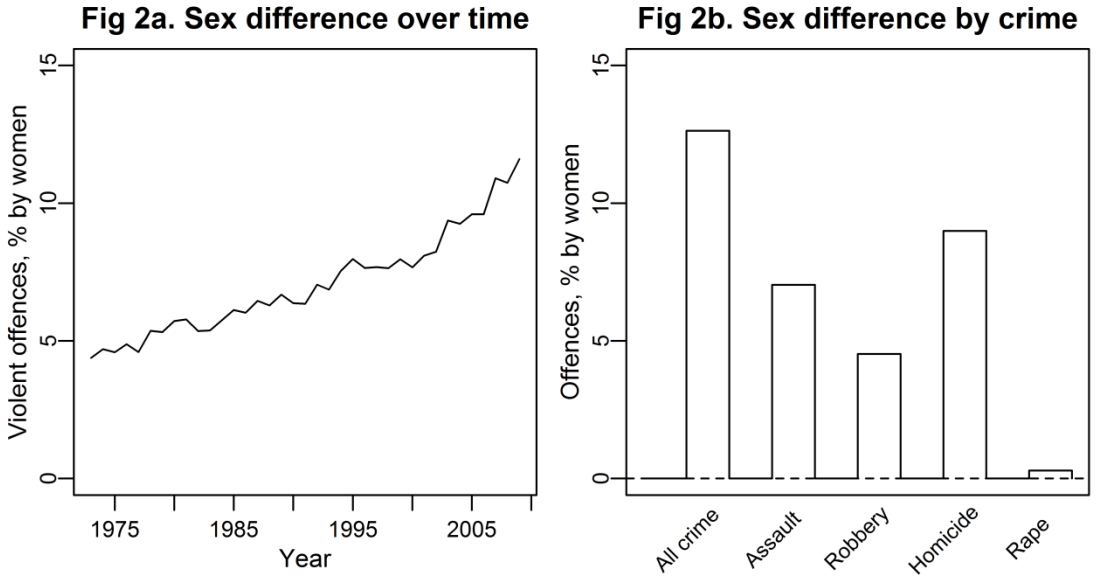


In Figure 1, the number of convictions for violent offences in Sweden in 2009 is plotted by age of the perpetrator. To take the population age distribution into account, the numbers are per 1000 individuals of that age, alive and living in Sweden in 2009. Since the age of criminal responsibility is 15 in Sweden, no perpetrators younger than this may be recorded in the Register of Criminal Convictions. As shown in Figure 1, the peak age for convictions of violent crime is 15-16, but the rate does not level of until age 30, after which it is quite stable until the late 40s.

Despite the stability of aggressive or antisocial behavior, many who commit crime or other serious antisocial acts eventually stop doing so. Indeed, for every individual, some antisocial act must be their last. Researchers have pointed out that this desistance from crime may be caused by some important event, or turning point, in the individual's life [57]. Research on turning points is of obvious interest, since it may identify interventions that could be applied to reduce criminal recidivism. Among the many turning points that have been suggested are entering military service, getting a good job, education, marriage, and becoming a parent [57, 58]. Though this has become something of a hot topic in recent years, current research has rarely been able to test competing hypotheses, or explain why each "turning point" is only associated with reduced antisocial behavior for some individuals.

1.1.4 Men are more violent than women

Men are more likely than women to be convicted of violent crime. In Sweden the life-time risk is about 10 times higher among men (e.g. Table 1 in Paper I). Though the exact magnitude varies, pronounced sex differences are evident for most antisocial behaviors and appears early in life. While there is no consistent evidence for differences in the first two years of life, from ages 4-5 boys are more aggressive and much more frequently diagnosed with CD than girls are [42, 59]. These differences persist over the life-course [49, 60].



The gender gap does seem to be universal, but the magnitude of the difference in both registered and self-reported crime has decreased during the past decades, at least in the United States and Canada [61], the UK [62], and Sweden [18]. Figure 2a shows the proportion of all convictions for a violent offence in the Register of Criminal Convictions where the perpetrator was female, increased from below 5% in 1973 to above 10% in 2009. Figure 2b shows how the proportion of all registered convictions varies with type of crime. Females account for about 13% of all convictions, some 7% of all convictions for assault, but less than 1% of all convictions for rape or sexual assault.

Given the relative rarity of violent crime among women, and the general tendency for research to focus on men, it is perhaps not surprising that few studies have focused specifically on risk factors for crime among females. Though this lack of data should be acknowledged, it seems that most risk factors for violent crime or antisocial behavior among men are also risk factors among women [63]. It is possible that men and women share the similar risk factors and liability to commit violent acts, but that being a woman acts as a strong protective factor [64]. In this view, women who, despite their gender, commit a violent offence would need to have been exposed to multiple or stronger risk factors than men. This is consistent with findings that female violent offenders seem to be at particularly high risk for mental health problems and traumatic life events [61, 63, 65, 66]. However, it also seems likely that some risk factors are gender specific. For instance, early menarche has been reported to be associated with antisocial behavior [67].

1.1.5 Antisocial behavior aggregates in families

Relatives of individuals with antisocial behavior are at an increased risk of also developing antisocial behavior [8, 9, 68]. Many studies have focused on the *intergenerational transmission* of antisocial behavior, i.e. on whether children of antisocial parents are more antisocial themselves. Intergenerational transmission has been reported for many types of antisocial behavior and related disorders, including criminal convictions [55, 69], violent offending [70], criminal careers [71], externalizing disorders [36], child abuse [72], aggression [50, 73], and partner violence [41]. The abundance of situations where the behavior of children mirror the problems of their parents have led some authors to speak of a general "intergenerational transfer of psychosocial risk" [74], while other focus instead on the specificity of the transmission, i.e. parental conduct disorder predicts conduct disorder in children more strongly than it predicts ADHD or anxiety disorders [75].

Familial aggregation of antisocial behavior would be expected from most theories of its development. It would for instance be expected if there are genes predisposing to antisocial behavior; if there are intergenerational continuities in socioeconomic status and socioeconomic status has an influence on criminal propensity, if children learn antisocial behavior from their parents, or if there are causal effects of bad parenting, abuse or neglect on later antisocial behavior [55].

Often, intergenerational transmission has been interpreted as evidence for the currently preferred of these hypotheses, despite inability to control for competing explanations. Influential in a Scandinavian context, Gustav Jonsson (1967) used the term ”social heritage” to describe that children are born into the social class of their parents [76]. Though most contemporary researchers would probably claim to acknowledge the importance of genes for the inheritance of traits, including psychological characteristics, it is striking how often studies of intergenerational transfer interpret their findings as support for a social learning perspective, with no control for genetic inheritance [41, 72, 73, 77]. For example, a 2004 review of the literature on the role of family-of-origin violence in men’s marital violence perpetration noted that almost none of the reviewed studies included a genetic perspective on the intergenerational transmission [78].

Many studies trying to explain the intergenerational transmission of antisocial behavior have focused on processes within the family, or factors shared by all family members. A recent review of quasi-experimental studies concluded that there is some support for an effect of harsh discipline, maltreatment, divorce, adolescent motherhood, parental psychopathology, and poverty [79]. However, familial aggregation of antisocial behavior would also be expected if there are individual-level causes of antisocial behavior that are themselves heritable or transmitted in families. For instance, severe mental disorders, substance abuse, and general cognitive ability have all been extensively studied as possible causes of crime, and they are all known to cluster quite strongly in families.

1.1.5.1 Severe mental disorder

The association of severe mental disorders, primarily schizophrenia and bipolar disorder, and violent crime has been subject to heated debate [80]. Case reports, and the experience of clinicians working in forensic psychiatry, have long suggested that an individual’s psychosis may be a contributing cause to violent criminal acts [80, 81]. Whether severe mental disorders are associated with an increased risk of committing violent acts on a population level has, despite this, been contested. The subject is sensitive since individuals suffering from schizophrenia or bipolar disorder are already at a disadvantage in society, and it has been argued that a focus on a risk increase for violence would only serve to increase stigma and discrimination [82]. In hindsight, this may have led to some overly careful interpretations of the available data. In a 1984 review, Mullen concluded that there was no proven association between severe mental illness and violent or other crime, though he also lamented the shortcomings of the extant literature [83]. Studies from the same era claimed that the association, if present, could be explained by confounding from age, sex, or socioeconomic status [80, 84]. Since then, a large number of population-representative epidemiological studies indicate that there is indeed a clear overrepresentation of violent crime among individuals with schizophrenia and/or bipolar disorder [85-94]. Meta-analyses have suggested that schizophrenia and bipolar disorder entail similarly

increased risks, four to five times higher rates of violence than in unaffected controls [87, 88, 95]. It is now also accepted that the association remains after adjusting for sex, age and demographic background variables [85-87, 91]. Whether the association is strong or not is a matter of opinion; the population attributable risk of violent crime has in Sweden been estimated to be 5% [96], and a review concluded that estimates of the population attributable risk consistently fall below 10% [94]. Current authors seem inclined to accept that there may be a causal effect of severe mental disorders on violence, and has moved to trying to test different theories of how the effect may be mediated. For instance, 20 years after his previously mentioned review, Mullen suggested that the effect of schizophrenia may be mediated by a host of factors; among them problems with social adjustment, educational failure or unemployment, and substance misuse [82].

1.1.5.2 Substance abuse

Substance abuse is a strong risk factor for aggression and violence [97, 98]. A Swedish register-based study reported that if the association of substance abuse disorder and violent crime were completely causal, removing substance abuse disorder would remove almost a quarter of all violent crime [99]. In support of a causal theory of the association of substance misuse and crime, studies have reported that substance misuse predict later antisocial behavior [100, 101]. It has also been observed that violent offenders are often under the influence of alcohol at the time of the offence. That alcohol may act as a *trigger* for violence has also gained support from case-crossover studies [102], and would fit with theories of the physiological effects of alcohol [103]. However, the effect may be reversed for benzodiazepines and cannabis, and the possible trigger effect of other substances remain uncertain [97, 98, 102]. The association of substance misuse and antisocial behavior is likely to be partly explained by confounding and reverse causality. There is strong evidence that delinquency predicts later alcohol and marijuana use [104], and that childhood CD or adolescent problem behavior predicts substance misuse [105-108]. Substance abuse and antisocial behavior share many risk factors that could confound the association, among them impulsivity [109], and low intelligence [110]. Given the strong correlation of substance use and antisocial behavior, it has been suggested that they are caused by a similar personality constellation or distinct realizations of a common latent phenotype, the externalizing spectrum [34, 93].

1.1.5.3 General cognitive ability

Many neuropsychological constructs have been reported as risk factors for antisocial behavior and delinquency, but the one with longest history is probably intelligence (see e.g. [23]). General intelligence or cognitive ability has been described as the ability to understand complex ideas, to adapt and learn quickly, to plan and to reason, and to solve problems by thinking [111, 112]. The definition of intelligence, and disagreements within the research field, are both closely tied

to the development of factor analysis, a group of statistical methods trying to explain observed correlations among a set of variables in terms of underlying, latent, factors [113, 114]. Briefly, Spearman developed a version of factor analysis in the early 1900s to address the high correlation among students' performance in different academic subjects, and concluded that one underlying factor, dubbed the *g*-factor, could explain these correlations [113, 114]. This finding fit well with the contemporary development of intelligence tests, where several tests, each representing slightly different cognitive tasks, were combined and the overall test result interpreted as a measure of general (non-test specific) ability. During most of the 20th century, developments in factor analysis were accompanied by an increasingly refined description of the intelligence construct, and on-going debate between intelligence researchers on the proper application and interpretation of factor analysis [113-115]. One influential perspective has been Cattel-Horns model, in which there was no *g*-factor but rather two factors representing fluid (Gf) and crystallized (Gc) intelligence, respectively [116]. Fluid intelligence was thought to involve adaptivity, learning, and innovation; while crystallized intelligence would entail using knowledge, verbal skills, and contextual comprehension. The model was later extended with several other "intelligences", among them processing speed, short-term memory, and visual processing [115]. Though the model explicitly did not contain a common factor influencing these aspects of intelligence, they were quite highly correlated to each other [114]. Today, the most prevalent description of intelligence is probably Carroll's hierarchical three-stratum model [113], sometimes combined with the Cattell-Horn terminology in the so-called CHC framework [115]. According to this view, a *g*-factor representing general intelligence is on the top level, and can account for as much as 50% of the variance in the individual intelligence tests. Below the *g*-factor are several second-level factors representing intelligences specific to groups of task or activities (corresponding to Catell-Horns Gc, Gf etc), and lowest-level factors representing cognitive abilities that are specific to each task or activity.

General intelligence, as measured by intelligence tests, is a strong predictor of school performance, income, job performance, and a wide range of socioeconomic outcomes [111]. Intelligence test results vary with age, increasing as individuals mature and learn, to eventually decline with advanced age [113]. Despite this, an individual's position relative to others of the same age is quite stable [111, 113]. The physiology of intelligence is largely unknown, but it is associated with brain volume (correlation 0.4, according to two reviews [117, 118]), and possibly with cortical thickness in specific brain regions [118]. Several neuroimaging studies suggest that intelligence is correlated to brain efficiency, measured as lower brain functional activity during moderately difficult cognitive tasks [119]. General intelligence is at least moderately heritable, with analyses on a sample combining six large twin cohorts estimating the heritability of high intelligence (being in the top 15%) to 50%, and the heritability of actual intelligence score to 55% [120].

The heritability seemed to be lower in childhood (41%), and higher in early adulthood (66%), while the estimated influence of shared environment changed in the opposite direction (dropping from 33% to 16%) [121].

This association of intelligence and crime has historically been subject to some controversy, in part due to different opinions on whether the general cognitive ability measured by intelligence tests represents a true entity, and in part due to a preference for structural risk factors in criminology. When Hirschi and Hindelang (1977) presented a literature review and asserted that there was a substantial association between intelligence and delinquency, they were purposefully provocative and claimed that the field of criminology had overlooked the association due to ideological blindfolds [122]. The response did not disappoint. Simons [123] called their hypothesis "neogenetic", asserted that "Experts in the area of intelligence no longer view IQ as a global mental ability which one inherits", and concluded that "Hirschi and Hindelang contribute little to the topic other than obfuscation. Indeed should their naïve view of IQ be taken seriously, the field of sociology would be taking a giant step backwards". In another reply, Menard and Morse (1984) estimated that individual characteristics in total explained less than 5% of the variation in delinquent behavior and concluded that "the IQ-delinquency hypothesis contribute nothing to existing delinquency theory" [124], a report that was in turn criticized by Harry and Minor (1986), who suggested that the conclusions of Menard and Morse were premature, and specifically questioned "their logic, sampling, analysis, and model specification" [125]. Gradually, this heat seems to have faded, probably with an increased acceptance of intelligence as a construct, and decreased influence of structuralist theories in criminology.

Today, it is widely recognized that there is an association between general cognitive ability and crime or delinquency. Indeed, Farrington counts low intelligence among the most important risk factors of offending [4]. The association has been found for both self-reported and officially recorded crime [126, 127], and appears stronger for repeat offenders and violent or more severe types of crime [48, 128-130]. Many studies of intelligence and crime has focused on men, but the association has been reported also among women [131-133] . It has been reported that *Verbal intelligence* would be stronger associated to antisocial behavior than *Performance intelligence*, as measured by the Wechsler intelligence scales. A meta-analysis concluded that the difference may be strongest in adolescence [134]. However, the question may be moot since the distinction between Verbal and Performance intelligence has been criticized for not being supported by factor analytical studies, and the difference may simply reflect that Verbal intelligence is a better measure of *g* [134].

Since the IQ-delinquency association has been reported to remain when controlling for sex, race, and childhood socioeconomic position [135-138], intelligence is often assumed to have a causal effect on criminal propensity,

potentially mediated by school adjustment/performance [122, 132, 139, 140]. In support of this hypothesis, low childhood intelligence predict later adolescent delinquency and adult violence [127, 141-143], and the IQ-delinquency association seem to be attenuated by adjusting for school performance variables [139, 140, 143]. As discussed in section 1.2.1.2, conditioning on a mediating variable often introduces bias in the estimates [144], so the mediation results should be interpreted with caution. Raising doubts on the importance of school performance, there is also a contemporaneous association of pre-school IQ and conduct disorder [133, 145]. Further, it should not be considered proven that there is a causal effect of intelligence on antisocial behavior. For instance, it has been argued that low intelligence may be a consequence of conduct disordered children's truancy and lack of education, or their decreased motivation or attention during IQ-testing [134].

Intelligence is a neuropsychological construct, the actual number based on factor analysis of a series of test results. Although it has been shown to be related to neuroanatomy and several other neurocognitive constructs, intelligence in itself does not necessarily have a manifestation in the physical world. Intelligence is associated to many similar constructs, and it is possible that one of these would better explain the association. For instance, studies have found similar associations with antisocial behavior for verbal intelligence as for executive functioning [128], and children and adolescents diagnosed with ADHD score on average 2-5 IQ-points lower than those without attention or hyperactivity problems [146].

1.2 CAUSAL INFERENCE

Although many correlates for crime and violence have been identified, the causation linking these variables is not well understood. With no clear understanding of the causal pathways, interventions and treatment programs may target the wrong factors, led astray by spurious associations, epiphenomena, or reversed causality. Without experiments, it is often impossible to actually test causal hypotheses, and it is well known that an observed correlation between two variables does not imply that one causes the other. However, every correlation *is* caused by something, and we must base our causal inferences on attempts to reason about what causal effects would best explain the associations we observe.

Formally, causal effects are often defined in terms of *counterfactuals*. Say a person was unexposed and we observe that person's outcome. The exposure is said to have a causal effect on the outcome when the outcome for the person would have been different if, *contrary to fact*, the person had been exposed. On a population level, we may say that X is a cause of Y, if there is anyone whose level of Y would change with a counterfactual change in X. This is of course impossible to observe, but experiments try to mimic counterfactual situations by holding all other conditions constant and intervening on the exposure. In a randomized controlled

trial, for instance, we would randomly assign individuals to different treatment conditions and compare the outcome in differently treated groups. Though individuals would differ in their intrinsic risk for the outcome, the randomization ideally ensures that these differences average out over the groups and only the difference in treatment remain as an explanation of any observed differences in outcome.

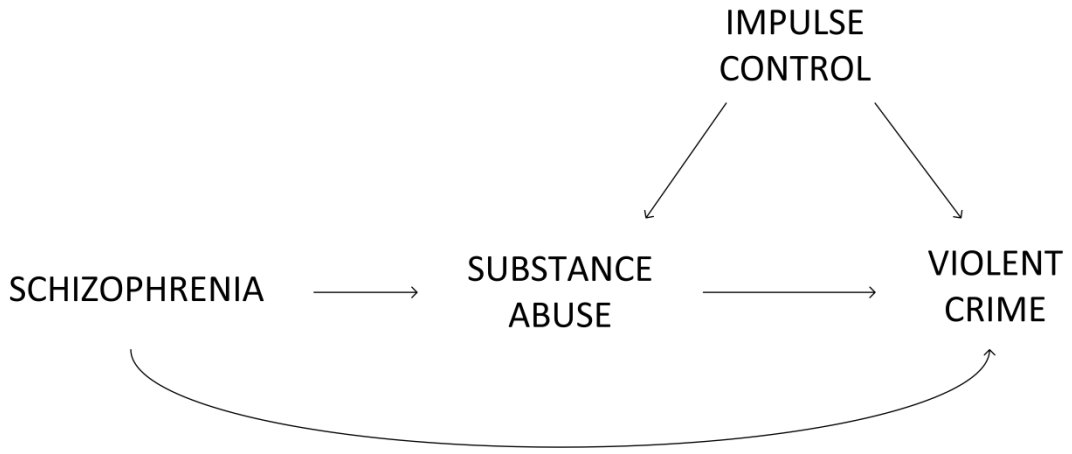
The difficulty with making causal inferences from studies of violence and crime is that they are predominantly observational rather than experimental [4]. In an observational study the researcher has not made any intervention, and simply observes associations as they appear naturally. Though a difference in outcome among differently exposed individuals may be due to a causal effect of exposure on outcome, it could also result from a causal effect of outcome on exposure (reverse causation), common causes of exposure and outcome (confounding), or from some selection of subjects who participate in our study. The challenge in observational studies is to, by appropriately taking these processes into account, produce comparisons that mimic counterfactual situations.

Naturally, in all studies, experimental or otherwise, a finding could also be due to random variation. Whether the observed difference is larger than expected by chance is *the* question in statistical inference, and in practice this is an important point. In this section, however, we are concerned with the question of how to interpret an association even if it has been measured with infinite precision.

1.2.1 Causal diagrams

Causal diagrams are an aid in discussing hypotheses of why a particular exposure is associated with an outcome. When formalized, they can be used for deciding what is needed to make causal inferences from a study. Directed acyclic graphs (DAGs), are one type of formalized causal diagrams that have been gaining popularity in epidemiology [147, 148], and is used in Papers II and III of this thesis. For illustration, Figure 3 is a simplistic DAG of the association of schizophrenia and violence. DAGs are directed, meaning that associations between variables are drawn as arrows pointing from one to the other. An arrow from X to Y should be read as "X may cause Y". DAGs are also acyclic, meaning that it should not be possible to follow arrows from X back to X. In other words, a variable may not cause itself. Since an arrow indicates the possibility of a causal association, the absence of an arrow indicates an assumption of no causal effect. In Figure 3, we are open to the possibility that schizophrenia influences substance abuse and violent crime, but we are assuming that schizophrenia does not influence impulse control. The variables in a DAG are connected by *paths*, along which arrows and other variables lie. Unless we have controlled for any variables, a path is blocked if it contains an *inverted fork*, i.e., if two arrows meet in a variable (→ X←), else it is open. In Figure 3, schizophrenia is connected to violent crime by three paths, *schizophrenia → violent crime*; *schizophrenia → substance abuse → violent crime*; and *schizophrenia → substance abuse ← impulse control → violent crime*. The first

Figure 3. Simplistic DAG of the schizophrenia-violent crime association



two paths are open, but the third is blocked by an inverted fork. Open paths make variables statistically dependent, and will contribute to an observed association between them. Conditioning on, i.e. controlling for, the variable in an inverted fork will unblock the path at that point. Conditioning on any variable in an open path will block it.

1.2.1.1 Confounders

Let us for the moment assume that Figure 3 is correct, and that we wish to assess the causal effect of substance abuse on violent crime. According to the DAG, substance abuse is linked to violent crime by three paths, *substance abuse→violent crime*, *substance abuse← schizophrenia →violent crime*, and *substance abuse← impulse control →violent crime*. The first path is the causal effect we wish to estimate, while the two other paths are through common causes of substance abuse and violent crime. All three paths are open, and will contribute to the crude association we would observe in our study. We say that this crude association is *confounded*, and that schizophrenia and impulse control are *confounders* of the substance abuse - violent crime association. The solution to confounding is to somehow control for the confounders, for instance through stratification, regression modeling, or propensity score matching. In DAG terminology, conditioning would block the paths through the variables, and only the causal path *substance abuse→violent crime* would be left.

1.2.1.2 Mediators

Let us keep assuming that Figure 3 is correct, and that we are interested in the association of schizophrenia and violent crime. As stated previously, the crude association will be a combination of two paths, one direct *schizophrenia →violent crime*, and one indirect *schizophrenia → substance abuse→ violent crime*. In the indirect path, the effect of schizophrenia is mediated by substance abuse, and substance abuse is called a *mediator*. Neither of these paths is due to some

confounder, we never go against the direction of an arrow, so the crude association would be a correct measure of the total causal effect of schizophrenia on violent crime.

Often, there is interest in testing how strongly an effect of X on Y is mediated by M. We might be tempted to adjust for substance abuse, closing the path *schizophrenia → substance abuse → violent crime*, and interpreting the remaining association as the direct effect of schizophrenia on crime, i.e. *schizophrenia → violent crime*. However, in Figure 3, conditioning on substance abuse would not only close the path *schizophrenia → substance abuse → violent crime*, it would also open the path *schizophrenia → substance abuse ← impulse control → violent crime*. This previously blocked path would contribute to the abuse-adjusted association, and it would not be a correct estimate of *schizophrenia → violent crime*. This phenomenon is sometimes referred to as *collider stratification bias* [144, 149, 150], and is often overlooked in studies attempting to perform mediation analysis.

1.2.1.3 Colliders

In the mediation example above, bias is introduced by adjusting for a variable in an inverted fork. Such variables are referred to as *colliders*. It may at first seem counterintuitive that statistical adjustment for a variable may introduce bias, but this is the same mechanism that leads to selection bias. Say selection of subjects into a study is influenced by high X and high Y. Individuals selected to be part of the study and have low X have increased likelihood to have high Y, for if they were not selected based on their high X, they must have been selected on something else. This will introduce a negative association of X and Y in the selected sample.

When we adjust for a variable, whether through regression modeling or otherwise, we are essentially estimating the association within strata of this variable. In the example of Figure 3, adjustment for the mediator indicates that we will look at the association of schizophrenia and violent crime among individuals *with* substance abuse separately from the association among individuals *without* substance abuse, and then combine these associations into one estimate. Individuals with substance abuse who do not have schizophrenia will have an increased likelihood of low impulse control, in turn increasing their likelihood of violent crime. Individuals without substance abuse who have schizophrenia will have a decreased likelihood of low impulse control, in turn decreasing their likelihood of violent crime. Stratification thus creates an inverse association between schizophrenia and violent crime. Controlling for substance abuse will yield an association that is a combination of to two open paths between schizophrenia and violent crime, working in opposite directions. The direct causal path which give a positive association between schizophrenia and violent crime, and the "backdoor path" through impulse control, which will give an inverse association between schizophrenia and violent crime. In Paper III, we show that

the within-pair estimate in sibling comparison studies suffer from a similar problem.

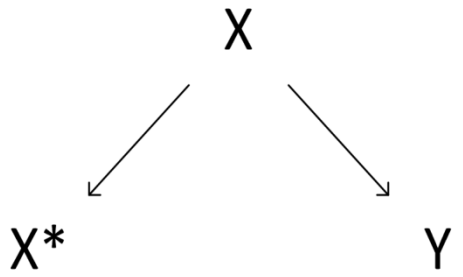
1.2.1.4 A note on path diagrams

Path diagrams, often used to illustrate structural equation models, are sometimes referred to as causal diagrams. These should not be confused with the DAGs described here, and used in Paper II and III. While DAGs are non-parametric visualizations of a causal scenario, path diagrams are fully parametric illustrations of linear equation systems. The paths in a path diagram represent linear regressions, and do not necessarily encode any causal assumptions.

1.2.2 Measurement error

It is unlikely that our measures completely correspond to the true factors we are trying to estimate. Trivially, our indicator variables may not actually represent the factors we are interested in, e.g. our interpretation may be wrong if we are measuring one subtype of cognitive ability but wish to make statements on full-scale IQ. But even when our measures are aptly named and interpreted, it is unlikely that we are measuring our variables with perfect precision. This imprecision may be called random measurement error and is illustrated with a DAG in Figure 4. If X* is the measurement of X, random measurement error will make the association of X* and Y weaker than the association of X and Y, yielding a bias towards the null. However, if the measurement error is not random with respect to the exposure, so that those observed as exposed would more often be incorrectly classified than those that are observed as unexposed, this *differential misclassification* may lead to bias in either direction.

Figure 4. DAG of measurement error



For continuous measures, the degree of measurement error may be expressed by the *reliability*, which is defined as the variance in true X divided by the variance in measured X. Since imprecision in the measurement would theoretically lead to an inflation of the variance, the reliability is a proportion between 0 and 1. The true reliability may be estimated by the test-retest reliability, repeating the same test several times on the same subjects. Since the subjects may be influenced by the previous test (e.g. by remembering the correct answers), test-retest reliability is not necessarily a valid measure of the degree of measurement error.

The degree of measurement error for a dichotomous exposure is often expressed in terms of sensitivity, the proportion among the truly exposed that are correctly

classified, and specificity, the proportion among the truly un-exposed that are correctly classified.

As shown in the DAG, the association of our measure, X*, and our outcome, Y, is due to their common cause, the true X. It is somewhat interesting that despite confounding being one of the biggest threats to causal inference, we are also capitalizing on it. Indeed, every studied association is confounded by the true causal factor our flawed indicators attempt to measure. If we were truly able to remove all confounding of an association between two imperfectly measured variables, there would not be any association left.

1.2.3 Using relatives as controls

Recall that causal inference would logically depend on comparing the observed outcome with what would have been observed under a different (counterfactual) exposure. Although we can almost never observe an individual under counterfactual exposure levels, we can try to approximate this by comparing "like with like". In other words, we may try to compare index and reference groups that are as similar as possible on all variables other than the exposure and outcome under study. This is what all statistical adjustments are trying to achieve, but it is perhaps most obvious in matched designs, where exposed-unexposed or case-control clusters are created based on specific matching variables. Like all statistical adjustments, this would only allow control for the covariates we have measured and to the degree that we have measured them correctly. As an attractive alternative, it has been suggested that we could use relatives as a reference group, automatically matching on everything the relatives share, potentially controlling for variables such as social disadvantage, genetics, and parenting, with no risk for measurement error.

Relatives has been used as controls for at least 70 years [23], but the practice has recently seen a resurgence in the behavioral sciences, in parallel to a focus on causal interpretation of risk factors, and the wide realization of the inferential weaknesses of observational research [7, 79, 151]. Monozygotic (MZ) twins, dizygotic (DZ) twins, ordinary siblings, and cousins (children of twins or ordinary siblings) have been most widely used as controls, but naturally any relative could be used. The comparison of relatives has alternatively been described as analysis of discordant twin pairs, as co-twin control studies, the children-of-twins design, sibship studies, between-within cluster modeling, family-fixed effects modeling, et cetera. Though analyzed slightly differently, all of these designs estimate the association of some exposure and outcome *within* relative-pairs, with the idea that such associations should be free from factors that are shared by the relatives.

Using relatives as controls is intuitively appealing, and the promise of potentially adjusting for unmeasured, and even unknown, confounding attracts many researchers, weary of the inherent limitations of observational studies. Perhaps because they seem so intuitive, they have received little attention from a

methodological perspective, and even texts recommending their use and advising on their interpretation do so in a heuristic manner [152-155], sometimes contradicting each other (cf [154, 156-158]). In a recent paper, we showed that sibling comparisons will produce estimates of the causal effect of exposure free from confounding by factors shared by the siblings, under the assumption of no other confounding and other biases [159]. The absence of non-shared confounders or other bias is rather unlikely in real-world applications, and economists have long been aware that under a linear model mean-reverting error and endogeneity due to omitted variables may lead to specific errors in the within-pair estimates [160, 161]. The language of economic research is quite different from that used in psychology and epidemiology, and it is quite possible that researchers in these fields have not recognizing that the econometricians refer to random measurement error and confounding. With the exception of McGue (2010), who acknowledged the influence of measurement error [7], the potential of sibling comparisons to increase rather than decrease bias seems to have gone mainly unrecognized in epidemiology and psychology.

1.3 GENETICS

1.3.1 Molecular genetics and genetic variation

To understand if and how genetic variation may contribute to variation in violent criminal behavior; it is helpful to have some understanding of what genetic variation is.

The human genome is made up by 22 autosomal chromosomes, the X and Y sex chromosomes, and the extranuclear mitochondrial genome. Together, these molecules consist of approximately 3 billion base-pairs, and contain 20-25,000 protein coding genes, and probably less than 1000 genes that code for untranslated but functional forms of RNA [162, 163]. Of the protein-coding genes, only 60 reside on the Y-chromosome and 13 in the mitochondrial genome. Regions commonly defined as genes make up ~25% of the human genome, but only ~1% are peptide coding regions [162, 163]. This reflects the large degree of "silent" or "junk" DNA that is interspersed both between genes, and within them. The regions between genes are mainly composed of repeated DNA motifs (transposons, tandem repeats and large duplications). The possible effect of these is relatively unknown, but repeated elements may influence the physical packing of the DNA strands, and thus influence expression of genes. Transposons are segments of DNA that may dislocate/duplicate and insert in new strands of DNA, and it has been suggested that this ability may give them a role in the evolution of the genome. The silent regions within genes are mainly introns, strands of DNA that are transcribed, but excised from the mRNA before translation.

Even in the fruit fly *Drosophila melanogaster,* probably the most studied multicellular organism in genetics, about half of the protein-coding genes are of unknown function. Roughly 20% code for enzymes, less than 10% for proteins involved in cell signaling and transportation, and 20% for transcription factors or

proteins involved in maintaining and replicating DNA [163]. Thus, much of the activity of the genome is dedicated to preserving the DNA molecules' integrity and function. Despite the obvious difference in structure and characteristics of various tissues, it has been estimated that about 10,000 of the 20-25,000 human genes are expressed in every human cell type [163]. This reflects how fundamentally important most of our genes are to life as we know it.

All genetic differences start out as mutations. Mutations are caused by damage to the DNA strand by radiation, chemical reactions, or errors during DNA replication or recombination. Mutations may take many forms, for instance a part of the DNA strand may be repeated, substituted, or deleted. Most large mutations are fatal, since so much of the genome is necessary for proper functioning of the organism. Most small mutations are completely neutral, since they will not influence genes or gene expression. Some mutations confer a positive advantage to their carriers, by making enzymes more efficient in the carrier's environment, or by increasing or decreasing the expression of some gene. Over generations, mutations will either disappear or spread through the population in a selection process influenced by chance and the mutation's contribution to the organism's fitness. Once a mutation is carried by 1% or more of the population it is commonly referred to as a *polymorphism*, and the alternative forms of the gene *alleles*. The most studied polymorphisms are Single Nucleotide Polymorphisms (SNP), which correspond to a substitution of one nucleotide with another, i.e. a change from one letter in the DNA code to another. The majority of polymorphisms among humans are expected to be completely neutral, but they may be used as markers to search for genetic variants with an effect on a particular phenotype.

Most of the genome, above 99%, is identical for all humans. Indeed, the sequence alignment overlap of human and chimpanzee genomes has been estimated at 98.63% and for human and gorilla genomes 98.25% [164]. It has been estimated that there are over 10 million SNPs in human populations, and in 2008, more than 1 million of these had been identified [163]. In comparison, it has been estimated that about 35 million single nucleotide changes separate humans and chimpanzees [165]. However, the difference between these species is also marked by many small insertions, deletions and inversions, as well as larger chromosomal rearrangements [166]. Regardless, the comparisons of ape genomes show that even slight genetic differences may cause striking differences in phenotype.

1.3.2 Quantitative genetics

The Mendelian laws of inheritance were recognized many decades before DNA was identified as the gene-carrying molecule. As early as 1918, Fischer used these laws to derive expected genetic correlations between relatives, and suggested that these may be used to estimate the proportion of the observed variance of a trait that was attributable to genetic variation in the population [167]. This theoretical proportion is today called the "heritability" of a trait, and remains a central concept in genetic epidemiology [168, 169]. The foundation for these calculations

is the assumption that the observed phenotype of an individual, *P*, can be viewed as the sum of the individual's genetic value (*G*, his or her genetic potential), and some deviation from this due to environmental influences (*Env*):

$$P = G + Env\,. \qquad \text{[Eq 1]}$$

The genetic and environmental values would here represent total effects of the individual's complete set of genes and environments. For a specific allele, say *B*, under the assumption of random mating, we may define the average effect of allele *B* as the average phenotypic deviation from the population mean among those who received allele *B* from one parent while all other alleles were received as a random sample from the population allele distribution. Summing the average effects of all alleles carried by an individual will give that *individual's additive genetic value*, A. The additive genetic value is sometimes equated with *the breeding value*, though the definitions do slightly differ. The breeding value is defined by mating an individual with randomly chosen mates and taking twice the offsprings' average phenotypic deviation from the population mean. An individual's breeding value is thus theoretically estimable in breeding experiments; while the additive genetic value is the individual's expected breeding value based on an additive model of theoretical population averaged estimates.

An individual's genetic value can be expressed as the sum of the additive genetic value, and deviations from this additive value due to statistical interactions between alleles at the same locus (the dominance deviation, D), or at different loci (often named epistasis, I), and we may write:

$$P = A + D + I + Env\,, \qquad \text{[Eq 2]}$$

and

$$\begin{aligned} Var(P) = {} & Var(A) + Var(D) + Var(I) + Var(Env) \\ & + 2Cov(A, Env) + 2Cov(D, Env) + \cdots\,. \end{aligned} \qquad \text{[Eq 3]}$$

The summation above would continue for all covariances of environment with genetic factors, and of genetic factors with each other. The above equations assume that there is no statistical interaction between genetic and non-genetic factors, i.e., that the additive effect of a gene is constant in all possible environments. This is unlikely to be true, and the formulae should be further extended by the set of gene by environment interactions:

$$P = A + D + I + Env + \{A + D + I\} * Env\,, \qquad \text{[Eq 4]}$$

and

$$Var(P) = Var(A) + Var(D) + Var(I) + Var(Env) + 2Cov(A, Env) + 2Cov(D, Env) + \cdots + Var(\{A + D + I\} * Env). \quad \text{[Eq 5]}$$

The heritability is defined as the proportion of the phenotypic variance that is due to additive genetic variation (*Var(A)/Var(P)*). In the absence of statistical interactions and covariance between genes and environments, this is the linear regression coefficient of breeding value on phenotype. As such, it would give an estimate of how much your phenotype says about your breeding potential. In the presence of covariance between genetic and non-genetic factors, or gene-environment interactions, the interpretation of the heritability may be more complicated.

1.3.2.1 Estimating the heritability

In practice, it is impossible to estimate the components of the phenotypic variance in Equation 5 since we lack measures of most of the genetic and environmental factors involved. Following Fischer's idea however, we may get rough estimates based on the covariance between relatives. The phenotypes of two relatives, {P_1, P_2}, will almost always be positively correlated since relatives partly share genetic and/or environmental causes of the phenotype. In the simplest, and most commonly used, quantitative genetic models, it is assumed that there is no covariance between genes and environments, no interactions either between genes, or between genes and environments. It is also assumed that the phenotype of relative 1 does not influence the phenotype of relative 2. Under such a model, we may set up the equation

$$Cov(P1, P2) = Cov(A_1, A_2) + Cov(Env_1, Env_2) = \rho_A Var(A) + \rho_{Env} Var(Env). \quad \text{[Eq 6]}$$

By combining relatives with different genetic and environmental correlations, such as full and half-siblings, we can construct an equation system and solve for the variance of A and *Env*. Dividing the estimated *Var(A)* by the estimated *Var(P),* we get an estimate of the heritability. While a value for ρ_A may be set based on knowledge of genetics and assortative mating, the ρ_{Env} does not have a theoretical basis. Rather than assigning a number for the environmental correlation, the environment is often split in two parts. Environment perfectly shared by relatives, C, with a correlation of 1 for relatives reared or living together, and "unique" or "residual" environment, E, with a correlation of 0 for all relatives. Since few, if any, factors will be completely uncorrelated in close relatives, most non-genetic factors should contribute to both environmental variances.

1.3.2.2 Heritability of dichotomous traits

The quantitative genetic theory described above assumes that the phenotype is a continuous, normally distributed variable. When the phenotype is binary, the formulae are extended through the liability-threshold model [169, 170].

According to this model, everyone has an unmeasured value of the liability to develop the phenotype, but only those having a liability score above a certain threshold actually do. Although the distribution of the liability is ususally unobservable, it is assumed to be an approximately smooth continuous distribution that could be transformed to a standardized normal distribution. Since the distribution is standardized, only the relative contribution, and not the exact values, of the variance components are interpretable. Statistically, tetrachoric correlations may be used to estimate relatives' correlation in liability [171], and mixed probit regression may be used to estimate heritability and other variance components of the liability [172].

1.3.2.3 Implications of model misspecifications

Comparing Eq 5 and Eq 6, it is clear that the model used to estimate the heritability has ignored many terms contributing to the phenotypic variance. We are also making additional assumptions regarding the genetic and environmental correlations of relatives. This has some implications for how results from quantitative genetic models should be interpreted [169, 173]. The pedigree-based heritability equations are like confirmatory factor analysis in that we set up a correlation structure, and given this structure, we estimate the factor loadings that would best reconstruct our observed correlations. Contributing variables with a correlation structure that does not correspond to one of the predefined factors will be split up, and load on two or more factors. Nature is, of course, indifferent to what we call these factors.

Say the true phenotypic variance is 50% caused by additive genetics, 15% caused by dominance deviations, and 35% caused by a set of environmental factors with overall twin correlation of 0.3. This would give phenotypic twin correlations of 0.755 in MZ twins, and 0.3925 in DZ twins. A classic ACE twin model would estimate that 72.5% of the phenotypic variance was due to A, 3% due to C, and 24.5% due to E. The estimates are correct for the correlation structure we set up, but the structure does not reflect the true causal structure, and we would be wrong to interpret A as additive genetics.

As shown in this simple example, fitting a classic twin ACE model in the presence of dominance deviations will overestimate the contributions of additive genetics while underestimating dominance deviations and environment, particularly the influence of environment highly correlated in twins. Non-additive epistatic gene effects will bias twin ACE estimates in the same direction as dominance deviations. Non-additive gene-environment interactions will load on A to the degree that the environmental factor is shared by twins and on E to the degree that it is not. If MZ twins share more similar environment than DZ twins, this will load on A. Gene-environment correlations are sometimes thought to reflect population stratification, or be caused by parents passing on both social heritage and DNA. In such instances, they would load on C. If the genetic correlation matrix was misspecified due to inbreeding or assortative mating, this would load on C,

and make A underestimate the influence of additive genetic effects. The impact of model misspecifications would be different for models based on other types of relatives, so we might trust results from quantitative genetic models more if we have seen a convergence of results across models based on different relatives.

1.3.2.4 Empirical estimates of heritability

Recently, methods have been proposed for estimating heritability directly from DNA similarities. One method focus on DNA shared by close relatives, where their proportion of the DNA shared *identical-by-descent* in the last generation is regressed on their phenotypic similarity [174]. This method has been used for studying human height; yielding heritability estimates similar to twin studies and supporting the idea that polygenetic variation across the entire genome contribute to phenotypic variation in height [175]. Alternative methods estimate the additive genetic correlation matrix of a population of unrelated individuals and use it in a mixed effects model of their phenotypes [176]. This has been done for a range of phenotypes, including intelligence [177], and BMI [178]. Though heritability estimates based on this method has so far been lower than corresponding estimates from twin models, the method is also known to underestimate the heritability [176]. These methods have not yet been widely used for behaviors, but they may offer an attractive, though more costly, alternative to the assumption-heavy pedigree-based methods in quantitative genetics.

1.3.3 Genetics of violence and crime

Although there have been few attempts to address violence or violent crime specifically, twin and adoption studies of other antisocial behaviors has estimated that about half the phenotypic variance could be attributed to genetic variation [20, 35]. Behaviors are complex traits, influenced by multiple contributing factors and we would not expect to find a single gene responsible for any behavior, including violent crime. It is assumed, instead, that many genes, hundreds or more, influence the trait through more or less indirect pathways [179]. For instance, if schizophrenia is indeed a contributing cause of violent crime, genes influencing the development of schizophrenia will also be genes for violent crime. With the completion of the Human Genome Project, some authors expressed great optimism about discovering the major genes involved in human psychology using genome-wide association studies (GWAS), where hundreds of thousands of genetic markers are simultaneously tested for association with a phenotype. Genome-wide studies have so far failed to find any replicable associations with CD or ASPD [180], but these studies were under-powered, the largest having fewer than 4000 subjects. With a growing realization of the power needed to identify specific genes in genome-wide studies, initial optimism has been replaced with more realistic expectations. Using a combined sample of 183,727 individuals to study human height, a highly heritable and easily measured phenotype, 180 loci, explaining about 10% of the phenotypic variance was identified [181]. Although

certainly more than nothing, this does not give us reason to hope that GWAS will yield great insight in antisocial behavior (which can reasonably be assumed to be a much more complex phenotype than height) in the near future.

An alternative to the non-hypothesis driven GWAS is to specifically test only a handful of candidate genes for association with a phenotype. Although with only weak effects, several genes encoding proteins involved in the serotonergic system and the stress response pathway are reportedly associated with aggression and antisocial behavior [182]. Results from candidate gene studies should be interpreted carefully, however, since initial findings have often proven difficult to replicate. The field is plagued by publication bias where null findings are much less likely to be published than positive findings, making even systematic reviews problematic.

Nevertheless, an interaction between the monoamine oxidase A (MAOA) gene and childhood maltreatment has been hailed as an example of both a specific gene for antisocial behavior (dubbed the "warrior-gene" by some media outlets and gene test vendors [183]), and evidence for the importance of gene-environment interactions. The MAOA gene encodes an enzyme that degrades monoamine neurotransmitters, such as dopamine, norepinephrine, and serotonin [184]. In 2002, a down-regulated allele of MAOA was reported to be associated with increased antisocial behavior among subjects exposed to severe childhood maltreatment, but possible with decreased antisocial behavior among subjects with no childhood maltreatment [185]. Interactions between the low-activity MAOA allele and adverse environment has been replicated several times, and withstood a meta-analysis [186], but the evidence is mixed (for a narrative review, see Gunter et al (2010) [180]). In a similar situation, where an interaction between serotonin transporter linked polymorphic region (5-HTTLPR) and stressful life events had been reported [187] and widely cited, a meta-analysis showed neither a significant interaction nor a significant association of the gene and depression [188]. A yet unpublished meta-analysis found no association of MAOA alleles and antisocial behavior [189], making it unlikely that the allele, even in the presence of an interaction with maltreatment, has a strong influence on antisocial behavior.

2 MATERIAL AND METHODS

2.1 SWEDISH REGISTERS

Papers I, II, and IV all involved data from a linkage of many of the Swedish nationwide registers. The principal linkage was completed in 2006 with follow-up ending at 2004-12-31, and was used in Papers I and II. The linkage was later extended and several new registers were added. This later linkage was used in Paper IV and enabled follow-up until 2009-12-31. The expanded linkage encompassed more than thirty registers, but only the following were used in the studies included in this thesis.

2.1.1 The Multi-Generation Register

The Multi-Generation Register identifies biological and adoptive parents of every person born 1932 or later, and registered as living in Sweden at any time since 1961. The register was constructed in 2000 by collecting data from Skatteverket's [the taxation office] census, and it has later been completed with information from other sources to increase coverage in older, now deceased, cohorts. Unless the biological/adoptive parents have actually lived in Sweden since 1947 (when the national personal identification number was introduced), it is not possible to identify them. In Figure 5, the proportion of all individuals with information on both biological parents is plotted against birth year, for Swedish-born and immigrants respectively. For individuals born in Sweden since 1968, the register has an almost perfect coverage, but it is less complete for older cohorts, adoptive children, and immigrants. As Figure 6 shows, adoptions in Sweden have undergone a dramatic change. Since the 1970s, adoptions of Swedish-born have

Fig 5. Coverage in the MGR

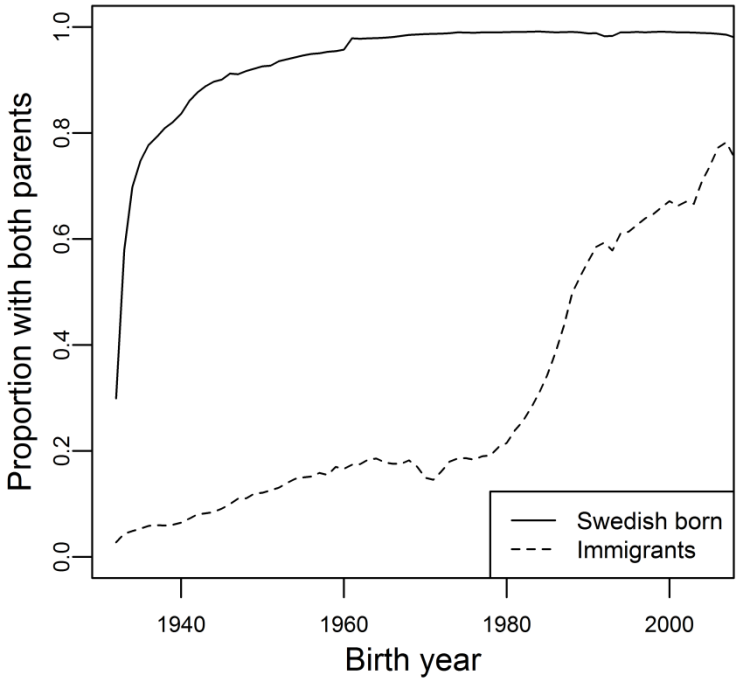
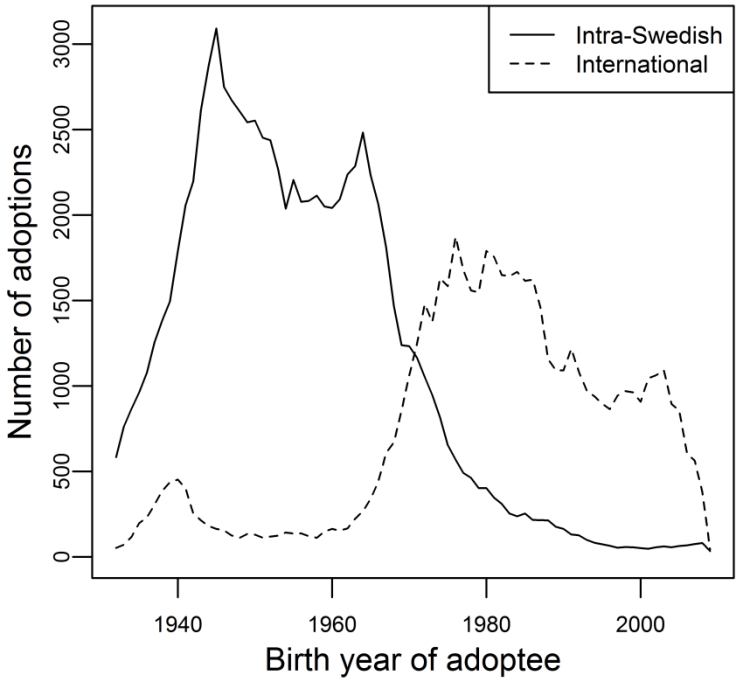


Fig 6. Trends in adoption

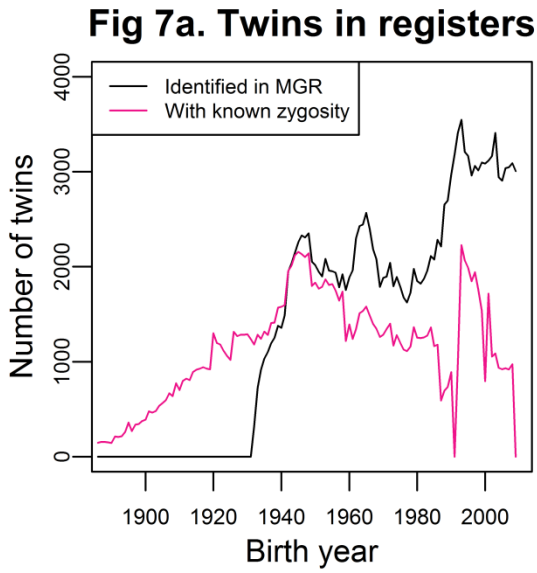


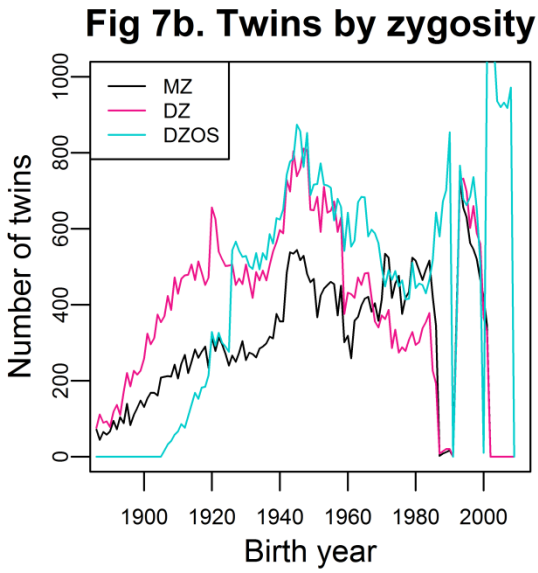
almost completely ceased, while international adoptions peaked around 1980, and has since decreased. Today, most of the intra-Swedish adoptions are adoptions by the partner of the mother or father [190].

With information on parents it is possible to create pedigrees for all individuals in Sweden. For instance, individuals sharing mother and father may be identified as siblings, and individuals sharing grandparents may be identified as cousins. Due to left-truncation of the data, relations that require information over three generations will be much less complete than those requiring only two.

2.1.2 The Swedish Twin Register

The Swedish Twin Register (STR) is a population-based register of Swedish twins born since 1886 and onwards. The register was established in the 1950s, when all parishes in Sweden was asked for information on all multiple births from 1886-1925. The twins thus identified were sent a series of questionnaires to determine zygosity and collect health information. Beginning in the 1970s, twins were identified through the birth register and several cohorts of twins have since been recruited. Notably, in 1998-2002 all twins in the STR born 1958 or earlier was invited to participate in computer-assisted telephone interviews, called the Screening Across the Lifespan Twin study (SALT), and in the 2006 all twins born 1959-1985 were invited to participate in a web-based questionnaire named the Study of Twin Adults: Genes and Environments (STAGE) [191]. These data collections contain many questions (in STAGE almost 1300 items) on life-style, health, life-events, and psychological traits. In this thesis, however, the STR was used only to determine zygosity of the twins used in Study II.





In SALT and STAGE, zygosity was determined from the following two questions: (1) "During childhood, were you and your twin partner as like as 'two peas in a pod' or not more alike than siblings in general?" and (2) "How often did strangers have difficulty in distinguishing between you and your twin partner when you were children?" Twin pairs who responded "alike as two peas in a pod" on the first question and "almost always" or "often" on the second were classified as MZ. If both twins responded "not alike" for the first question and "seldom", "almost never", and "never" for the second, they were classified as DZ, and other twin pairs were classified as "not determined". Opposite sexed twin pairs were, of course, always classified as DZ. In the SALT sample, a validation by genotyping showed that this algorithm misclassified only 2 of 199 twin pairs [192]. The participation rate was high in SALT, but substantially lower in STAGE. Since, aside from opposite-sexed DZ twins, only twins participating in studies can have their zygosity determined; this could introduce some selection bias in twin studies. Figure 7 shows the number of twins identified at birth and with known zygosity in the twin register (Figure 7a), and the number of twins of each zygosity (Figure 7b). In the period 1925-1958, the coverage of the STR is excellent, and there is an almost identical amount of same-sex and opposite-sex DZ twin pairs. This corresponds to the SALT data. After 1958, there is a growing disparity between the total number of twin births, and the number with known zygosity. The number of known opposite-sexed DZ makes a smooth transition from SALT to STAGE, and does not change dramatically until about 1980 (with the advent of in vitro fertilization). Among same-sexed twins, however, the lines cross over and in the period of about 1970-1985 there are more known MZ than DZ twins.

2.1.3 The Register of Criminal Convictions

The register of criminal convictions, or the crime register, contains information such as offence, date, and sentence for all convictions in Swedish lower courts (*tingsrätt*) since 1973 and onwards. It does not contain information on possible

Fig 8. Violent offending by birth year

changes in verdict (*dom*) or sentence (*påföljd*) after appeal to higher courts. Although the rate of appeals has increased for criminal code violations (from 7% in 1975 to 10% in 1993), the rate of substantial changes in higher court decreased during the same time (19% in 1975, 8% in 1993) [193, 194]. Thus, the rate of misclassification in our data due to changes in conviction status after appeal should be fairly constant at about 1%. The register information does not cover the circumstances of the crime, such as the identity of the victim, unless this is somehow captured by the paragraph invoked by the verdict. Further, crimes are registered regardless of medicolegal insanity at the time of perpetration, even if this leads to a sentence of forensic psychiatry treatment, and the Swedish system does not allow plea-bargaining. In Sweden, the age of criminal responsibility is 15 so crimes committed before this age are not registered. As shown in Figure 8, these truncations cause the proportion ever convicted of a violent crime to vary with birth year. Individuals born early are only at risk to enter the register when they are older, past the peak age of violent offending (cf Figure 1). Individuals born later have reduced time-at-risk, reducing their likelihood to have been convicted yet.

2.1.4 The Swedish Military Service Conscription Register

For more than a century, between 1901 and 2010, Sweden's military was based on a system of national conscription. Though the proportion who actually participated in military service dropped markedly during the 1990s, enlistment was mandatory for all male Swedish citizens until 2007. In the 1990s, less than 5% did not enlist, usually due to somatic illness or intellectual disabilities [114]. At the time of enlistment, which took place at approximately 18 years of age, conscripts underwent a medical examination and a battery of physical and

psychological tests. From the 1950s and onward, these tests were recorded in the Swedish Military Service Conscription Register.

The first intelligence test intended specifically for use in the Swedish conscription was the Swedish Enlistment Battery 1944 (SEB44), using the US "The general classification test" as a model, and was influenced by Spearman's idea of a general *g*-factor [114]. This general aptitude was thought useful for assigning conscripts to appropriately demanding service position. The test was changed several times over the years (major changes were made in 1948, 1949, 1954, 1959, 1967, 1980, and 1994) [114]. Though the earliest tests were developed with focus on psychometric properties and explicitly aimed to estimate general cognitive ability, the focus gradually shifted to estimating skills and aptitudes that would be most useful in the military. In the SEB67, a third of the items addressed technological aptitude, and factor analysis had been abandoned for a simpler summation score [114]. With the SEB80, the test used in Paper IV, psychometric considerations started to make a return to the Enlistment Battery. Compared to the SEB67, the SEB80 supposedly had better reliability, and aimed to better estimate general cognitive ability [114, 195]. The SEB80 consisted of four subscales with 40 items each, originally aimed at capturing different aspects of cognitive ability (verbal, spatial, inductive, and technological). However, validation studies showed that, while the overall tests score was a good measure of general cognitive ability (*g*) or fluid intelligence (Gf), the test could not reliably estimate lower order intelligence factors [196]. SEB80 was replaced in 1994, with a new computer-assisted test better suited at estimating also crystallized (Gc) and so-called general visualisation (Gv) intelligence [197].

In 2009, the Swedish Riksdag voted to replace national conscription with a voluntary standing army, but the general enlistment tests had been removed already in 2007.

2.1.5 The National Censuses

In the period 1960-1990, population and housing censuses were performed in Sweden every five years. In our first linkage, I had access to the censuses of 1960, -70, -80, and -90. In the extended linkage I also had access to the 1975 and -85 censuses. The censuses combined questionnaires with information obtained from various registers (e.g. of taxed income). Though the exact items changed during this time, the aim was to capture both financial and social aspects of every Swedish citizen's life, and the censuses contain information on occupation, income, "position in household", and type and size of habitation. In studies I and IV, we used the censuses to obtain information on childhood socioeconomic variables. In studies II and IV, we used them to differentiate siblings registered as living in the same or different households in childhood.

2.1.6 Migration data

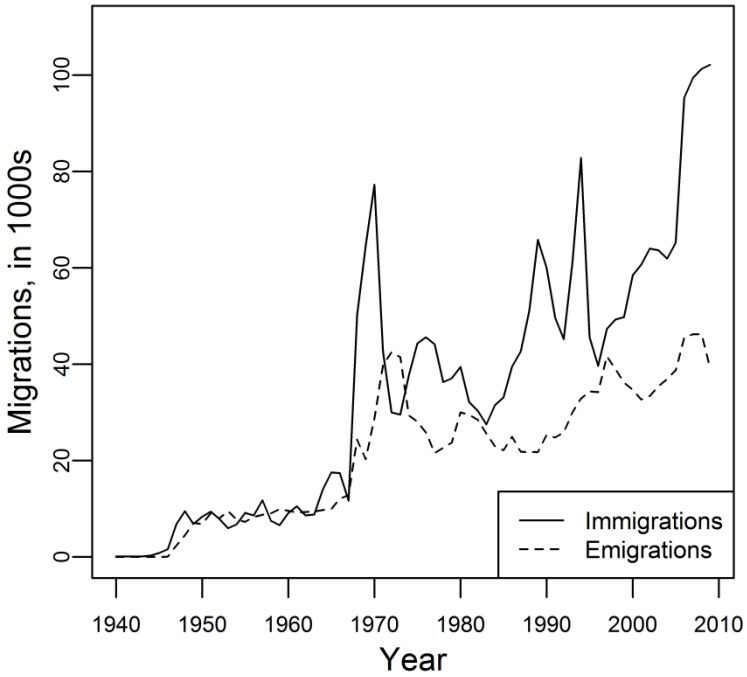
The Total Population Register contains information on country of birth and dates of all immigration (from 1969) and emigration (from 1961) events. In our linkage, birth countries had been aggregated into several large world regions (e.g. "Asia"), but the information was detailed enough to separate Swedish-born from Scandinavian-born, and from individuals born in other countries. Immigration and emigration data were used to censor individuals in Study I, and to assess the potential impact of censoring in Study IV. Sweden has a relatively high rate of immigration, and about 20% of young adults in Sweden are currently first generation immigrants. However, Sweden also has quite high rate of emigrations. Figure 9 shows the number of migration events by calendar year.

2.1.7 The Cause of Death Register

The Cause of Death Register contain information, including ICD-coded causes of death, on every deceased individual who were registered as living in Sweden at the time of death. It does not contain information on deaths among immigrants who had not received a permit of residence before their death, or deaths among tourists or visitors. It does, however, contain information on deaths among Swedish citizens who died abroad. The Cause of Death Register has information since 1952, but is considered complete since 1961.

In the studies in this thesis, the Cause of Death Register was used to get information on censoring, i.e. the death date for all individuals. Since our register linkage included parents of individuals born as early as 1932, and the Cause of Death Register was started in 1952, older generations may include deceased individuals who are seemingly alive. However, since index individuals had to be alive in 1961 to be included in the Multi-Generation Register, this should not be a

Fig 9. Trends in migrations



problem in the index generation.

2.2 STATISTICS

2.2.1 The nested case-control study

In study I, we performed a *nested case-control* study to estimate familial risks for violent offending. Generally speaking, a nested case-control study is defined as a case-control study that is nested within a specified cohort, and where the controls for each case are sampled from the individuals who were at risk of becoming cases at the event time of the case's outcome [198, 199]. An individual may appear several times as a control, and, if he/she later develops the outcome, as a case. This type of sampling is sometimes referred to as *risk set sampling*, since it samples from the so called risk set defined by the case, or as *density sampling*. According to the influential text book *Modern Epidemiology* by Rothman et al. [200], the latter name refers to that the sampling enables estimation of incidence rates, which are also known as incidence densities. Nested case-control designs are prospective designs, and may be considered when information on exposure or some confounder is expensive to collect, or as a way of reducing computational time while still retaining most of the power in the analysis.

While Rothman argues that all case-control studies should be seen as nested case-control studies, case-control designs may be performed in cross-sectional settings, or the design may prevent present cases to be past controls. Compared to these designs, the nested case-control has several advantages. Since at least some information is known for the larger cohort, absolute risks may be calculated through the known sampling probabilities. Further, the risk set sampling means that odds ratios in a nested case-control are ratios of odds of incidence, not prevalence. It can be shown that when the time intervals where controls are matched to cases are infinitesimal, the likelihood of a proportional odds model in the risk set sampled data will be identical to the likelihood of a proportional hazards model[198]. Except for sampling variability, and assuming that we have not matched on any covariates, the estimated incidence odds ratio of a nested case-control study should thus be identical to an estimated incidence rate ratio in the full cohort.

In *Modern Epidemiology*, Rothman does not mention sampling frames, but states instead that controls should be selected so the exposure distribution is the same among controls, as it is in the source population of the cases[200]. He further argues that as long as this is achieved, odds ratios from case-control studies should be reported as hazard ratios, and that the *rare disease assumption* often invoked as a motivation for interpreting odds ratios as risk ratios, is not needed. However, since at least one case is always removed from the source population before the controls are sampled, the controls cannot have exactly the same exposure distribution as the source population of the cases, unless the sampling frame is infinitesimally small. As the sampling frame grows wider, and more cases

are removed from the source population before controls are selected, the incidence odds ratio will move further from the hazard ratio [198]. Unless the outcome has a very high incidence, the incidence odds ratio will still be a very good approximation of the hazard ratio. We may not need a "rare disease assumption" to interpret the OR from a nested case-control study as a hazard ratio, but we do need a "not very common disease assumption". To avoid making any such assumptions, we present the results in Paper I as ORs rather than hazard ratios.

2.2.2 Clustered data and GEE

Many commonly used statistical techniques, such as the family of Generalized Linear Models (GLMs), assume that individual observations are independent, conditional on the modeled covariates. Whether we are using family data to estimate familial risks or wish to control for shared confounding, data sets based on families will violate this assumption. Since we are sampling clusters of similar individuals, each additional individual may not contribute the same amount of information as if everyone in the sample were independent. If not taken into account, this may lead to erroneous confidence intervals and p-values.

Obviously familial aggregation is not necessarily a problem; indeed, estimating this dependence of relatives is the very aim of heritability studies. When clustering is a nuisance, e.g. when estimating unpaired associations in the sibling comparison in Paper IV, we may account for it by bootstrapping, or less computationally demanding, by calculating "robust" standard errors. This is often done with a type of generalized estimating equation (GEE) also referred to as an *independence working model* [201]. This model is a multivariate extension of the score equation of the corresponding GLM, using the Huberized estimator of the sample covariance. Let $\mathbf{y_i}$ be the vector of observed outcomes for family i, and $\boldsymbol{\mu_i(\beta)}$ be the modeled response vector for family i (a function of the vector of regression coefficients), and $\mathbf{V_i}$ be the covariance matrix of outcomes in family i. Under the independence working model, we set $\mathbf{V_i}$ to a diagonal matrix, treating the outcomes within each cluster as independent. The score equation to be solved is then

$$S(\boldsymbol{\beta}) = \sum_{i=1}^{n} \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}\right)^T \boldsymbol{V}_i^{-1}\{\boldsymbol{y}_i - \boldsymbol{\mu}_i(\boldsymbol{\beta})\} = 0\,. \qquad \text{[Eq 7]}$$

The variance of the regression coefficients are given by the "sandwich" formula $\boldsymbol{H}_1^{-1}(\boldsymbol{\beta})\boldsymbol{H}_2(\boldsymbol{\beta})\boldsymbol{H}_1^{-1}(\boldsymbol{\beta})$ where $\boldsymbol{H_1}$ (the bread of the sandwich) is the Hessian, the standard covariance matrix estimator, and the "meat" of the sandwich is

$$\boldsymbol{H}_2 = \sum_{i=1}^{n} \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}\right)^T \boldsymbol{V}_i^{-1}\{\boldsymbol{y}_i - \boldsymbol{\mu}_i(\boldsymbol{\beta})\} \times \{\boldsymbol{y}_i - \boldsymbol{\mu}_i(\boldsymbol{\beta})\}^T \boldsymbol{V}_i^{-1} \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}}\right). \qquad \text{[Eq 8]}$$

In more general applications of GEE, the covariance matrix $\mathbf{V}_i$ is often estimated from data rather than set to be independent. In analysis of longitudinal data for instance, the fit of several different covariance structures (e.g. autoregressive, exchangeable, or Toeplitz) might be compared before deciding which is most appropriate for the data at hand [202]. However, assigning some covariance structure other than independent would in general be expected to change the estimated regression coefficients (if only slightly) [201]. Heuristically, this change would come from information borrowed from the other observations in the same cluster. In the context of family-clustered data, however, we have no reason to believe that our estimated regression coefficients are wrong; we simply wish to get more appropriate confidence limits. It seems clear that, despite recommendations to the contrary [153], these other covariance specifications should be avoided in these situations.

2.2.3 Probit GLMM

In Paper II, we estimated the heritability of violent offending by using a generalized linear mixed model (GLMM) with a probit link function. In general, a GLMM may be written as

$$g\{E(y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{z}_{ij})\} = \boldsymbol{x}_{ij}{}^T\boldsymbol{\beta} + \boldsymbol{z}_{ij}{}^T\boldsymbol{b}_i\,, \qquad \text{[Eq 9]}$$

where y_{ij} is the outcome and $\boldsymbol{x}_{ij}$ is the fixed covariates of the jth individual in the ith cluster (in our applications, the clusters will be families). The fixed effects are described by the regression coefficients, $\boldsymbol{\beta}$. The $\boldsymbol{z}_{ij}$ is a vector of known correlations, describing how the random parameters $\boldsymbol{b}_i$ are shared by members of the cluster. The link function, $g\{\}$, may define, for instance, linear, or logistic regression. By using the probit link, the inverse standard normal distribution function, and assuming that all random effects followed a normal distribution, we modelled the binary outcome as coming from a standard normal distribution with a distinct threshold. While everyone is assumed to have a value on this underlying liability, and this liability has some known correlation pattern in families, we only observe the outcome for individuals with liability higher than a certain score. This corresponds to the liability-threshold model described in section 1.3.2.2.

The likelihood for one cluster under the mixed probit regression model can be shown to be an integral over the multivariate normal distribution with covariance described by $\boldsymbol{z}_{ij}{}^T\boldsymbol{b}_i$ [172]. Where $y_{ij}=1$, the integration is taken from negative infinity to ,$\boldsymbol{x}_{ij}{}^T\boldsymbol{\beta}$ and where $y_{ij}=0$ from $\boldsymbol{x}_{ij}{}^T\boldsymbol{\beta}$ to infinity. To compute this integral, we used an algorithm by Genz based on Cholesky decomposition and a Monte Carlo approximation [203]. We obtained maximum likelihood estimates (MLEs) by optimizing the sum of individual log-likelihoods using the function *optim()* in R.

Approximate 95% confidence intervals were estimated from the profile log-likelihood of each variance component. In other words, the log-likelihood was

calculated when the specific variance component was fixed to a certain value, and the other variance components were set to the MLE they would have at this value. We fitted a smoothed function through log-likelihoods over a range of parameter values, and could read out confidence limits where this function exceeded a threshold corresponding to the requested alpha level [204]. The confidence intervals are approximate since we did not re-estimate the fixed components for each parameter value. This saved a lot of computation time, and the estimates of the fixed effects should be largely independent of the estimates of the variance components.

2.2.4 The between-within model

The between-within model is a flexible framework for performing sibling comparisons. The model is a simple extension of an ordinary GLM, where the individual's outcome is modelled as a function of the individual's exposure and the pair's mean exposure. If Y_{ij} and X_{ij} is the outcome and exposure for individual j in sibling pair i, the between-within model would be

$$g\{E(Y_{ij}|X_{ij}, \overline{X}_i)\} = \alpha_{BW} + \beta_W X_{ij} + \beta_B \overline{X}_i \ , \qquad \text{[Eq 10]}$$

where the expected values of Y, conditional on the covariates, follow some defined distribution, e.g. the normal or binomial distribution. The link function, *g{}*, puts the sibling comparison in a generalized linear model framework. Different link functions enable for example linear, logistic, probit, and log-linear regression. The exposure-outcome association is divided into a within-pair effect β_W, and a between-pair effect β_B. The β_B, though by some authors considered of great interest [153], is usually considered non-informative and is not interpreted. Despite some trepidations [152], the between-within model has the same form and interpretation for dichotomous exposures, though it may be recommended that the sibling average is then modeled as a categorical variable, to avoid additional parametric assumptions.

2.2.4.1 Comparison to other models

When the exposure is dichotomous, the β_W from a between-within model will give the same result as an analysis restricted to exposure discordant pairs [205]. When the outcome is dichotomous, sibling comparisons in psychology and epidemiology are often done using conditional logistic regression. This is similar to fixed-effects models often used (also for continuous outcomes) in econometrics, where each sibling pair is modeled with a pair-specific intercept. Though similar, this is not the same model as the between-within model, which conditions on the pair-mean in exposure. For general link functions, it has been shown that the within-estimate from a between-within model will be the marginal (non-pair specific) estimate standardized to the confounder distribution among the exposure discordant pairs, while the conditional estimate from fixed-effects models will be a weighted average of the pair-specific associations (e.g. ORs) [159]. For the linear link

function, these will coincide, but in general the models will not give quite the same estimate.

Sibling comparisons of time-to-event outcomes may be done in a stratified Cox regression [206]. Although it is also possible to directly use between-within models for survival data, the potential difference in efficiency and robustness of these models has not yet been addressed.

2.2.5 A note on non-collapsibility

Logistic and probit regression models are sometimes referred to as "non-collapsible", a property with some implications for how results from different samples, or adjusted for different covariates should be compared [207]. Suppose we are fitting a GLM to the association of X and Y as

$$g\{E(Y|X = x)\} = \alpha + \beta x\,, \qquad \text{[Eq 11]}$$

and compare this model with another linear model where Y further depends on C

$$g\{E(Y|X = x, C = c)\} = \alpha^* + \beta^* x + \gamma^* c\,. \qquad \text{[Eq 12]}$$

For non-collapsible link functions, $\beta \neq \beta^*$ as long as C is associated to Y, even if there is no association of C and X. In other words, the regression coefficient for X will be changed by including C in the model, even when C is not a confounder of the Y-X association. In epidemiology, non-collapsibility is a well-known property of the OR, but its implications for the interpretation of results from logistic regression is often not acknowledged. In short, conditioning on a C (e.g. by stratifying, or by including in a regression model) which is associated with Y, but not with X, will increase the regression coefficient for X [207]. This has implications for the comparison of marginal, population averaged estimates, and conditional, cluster-specific estimates [159, 205]. This includes co-twin control and other sibling comparison designs, though the importance of non-collapsibility in the comparison of within-pair and unpaired estimates has, to my knowledge, not been acknowledged in the literature.

3 STUDY SUMMARIES

3.1 PAPER I – VIOLENT CRIME RUNS IN FAMILIES

In Paper I, we attempted to quantify the familial aggregation of violent offending. We included every individual identified in the Multi-Generation Register born between 1900 and 1989, and defined the outcome as the first criminal conviction for a violent offence during 1973-2004. The study estimated familial risks among full siblings, maternal and paternal half-siblings, adoptive siblings, biological and adoptive parents, grandparents, aunts and uncles, cousins, and spouses (defined as a person with whom one had had one or more children). For all these relative types, we constructed all possible pairs, or dyads. In some cases, e.g. for cousins, the number of dyads far exceeded the number of individuals in the study. We then performed a nested case-control study where we matched on the index person's birth year, sex, world region of birth, and the relative's birth year and sex. If the relative had any conviction for a violent crime 1973-2004, the index person was considered "exposed", and the difference in exposure was compared between case and control individuals using conditional logistic regression and standard errors based on a robust covariance matrix estimator [208].

3.1.1 Results

Table 1 summarizes the main findings of the paper. Overall, we found significantly increased risks for violent crime among all studied types of relatives. The risks were highest among close relatives, and declined with decreasing genetic and environmental relatedness. Stratifying by gender revealed a great difference in OR between male-male and female-female relations. This was perhaps not completely unexpected given the large gender difference in violent offending. Note that adoptive relatives are in some situations also biologically related, since a child may be adopted by e.g. an aunt or grandmother. The highest male-female risk increase was for spouses, individuals who had had one or more children together, indicating strong assortative mating for violent criminality. We also found that the familial risks on the OR scale were modified by childhood socioeconomic status, age at first criminal conviction, and subtype of violent offending.

3.1.2 Considerations

The exposure in a nested case-control study is often defined as the exposure status/history at the event time of the outcome, so no "future" information would be used [200]. As often when studying familial risk, we did not consider the relative's criminal conviction as causing a conviction for the index person, we simply wished to estimate the familial clustering. For this purpose, we would have preferred to have total life-time conviction data, but lacking this we used the relative's total time-at-risk.

Table 1. Familial risk for violent crime in the Swedish total population 1973-2004

Relation to index person	Familial risk: Odds Ratio (95% CI)			
First degree relatives	Overall	Male-Male	Female-Female	Female-Male
Parent	3.5 (3.5-3.6)	3.3 (3.3-3.4)	6.3 (5.7-6.9)	4.3 (4.1-4.5)
Sibling	4.3 (4.2-4.3)	4.2 (4.1-4.3)	8.1 (7.4-9.0)	4.4 (4.2-4.6)
Second degree relatives				
Grandparent	2.0 (1.9-2.0)	1.8 (1.8-1.9)	3.1 (2.4-4.0)	2.2 (1.9-2.4)
Aunt or uncle	2.3 (2.3-2.3)	2.2 (2.2-2.3)	3.2 (2.8-3.6)	2.6 (2.5-2.7)
Maternal halfsibling	2.1 (2.1-2.2)	2.1 (2.0-2.1)	3.0 (2.6-3.5)	2.4 (2.3-2.6)
Paternal halfsibling	1.7 (1.7-1.8)	1.7 (1.6-1.8)	2.0 (1.6-2.4)	1.8 (1.7-2.0)
Third degree relatives				
Cousin	1.9 (1.9-1.9)	1.9 (1.8-1.9)	2.2 (2.0-2.4)	2.0 (1.9-2.1)
Unrelated				
Spouse	5.2 (5.1-5.3)	NA	NA	5.7 (5.6-5.9)
Adoptive relations				
Adopted child	1.5 (1.2-1.9)	1.4 (1.1-1.9)	10.0 (1.3-79.4)	1.8 (0.7-4.4)
Adopted away child	1.9 (1.7-2.1)	1.8 (1.5-2.0)	6.5 (2.4-17.2)	1.8 (1.2-2.7)
Adopted sibling	1.1 (1.0-1.3)	1.0 (0.8-1.2)	3.5 (1.4-8.8)	1.5 (0.9-2.4)
Adopted apart sibling	1.7 (1.3-2.1)	1.5 (1.2-2.0)	1.1 (0.2-4.9)	2.4 (1.4-4.2)

In this study, we nested the case-control study in the cohort defined by the Swedish total population registers. We already had exposure and covariate information on the complete cohort. So why perform a nested case-control study in the first place? First, by keeping all observations convicted of violent crime, but only a sample of the non-convicted, we reduced the data considerably but kept most of the information [199]. This led to swifter analysis, while only slightly reducing the precision of the estimated familial risks. Second, matching on birth year (and survival time) enabled adjustment for these factors without considering a specific parameterization in the regression model. It was our hope that this matching would also safe-guard against attenuation of the familial risks due to left truncation [209], i.e. misclassification since no conviction information were available before the register start-up in 1973. On the downside, the matched case-control data, while convenient for estimating the incidence OR, makes it difficult to estimate absolute risks of violent crime.

At the request of a reviewer, we complemented the ORs with tetrachoric correlations. The tetrachoric correlation of two dichotomous variables is calculated by assuming that variables are both generated from normal distributions with distinct thresholds [171]. If an individual is positive on the dichotomous variable, this indicates that his value on the underlying continuous variable is above the threshold. The correlation in the underlying normal distributions would be the tetrachoric correlation. However, the calculation does not adjust for the matched data structure, and the case-control sampling means that the assumption of an underlying normal distribution is almost certainly violated. If the assumption of an underlying normal distribution holds in the whole population, then in our case-control sample we have retained all

observations that scored above the threshold, but only a subsample of those that scored below the threshold. Clearly the resulting distribution would not be normal, and the tetrachoric correlations reported in the paper should be interpreted with caution.

3.2 PAPER II – HERITABILITY OF VIOLENT CRIME

In Paper II, we attempted to estimate the relative importance of genetic and environmental factors on the propensity of violent offending. We identified twins, full siblings, maternal and paternal half-siblings, adoptive siblings and adoptive parents among all individuals born in Sweden in 1932-1988. The outcome was defined as any conviction for a violent crime in 1973-2004. Variance components were calculated with probit GLMM. Since we knew from Paper I that there was assortative mating for violent crime, we performed sensitivity analysis with variance components estimated as a function of the genetic correlation of spouses. For this purpose, we needed to derive the expected additive genetic correlation among half–siblings. We also needed to estimate a likely range for the genetic correlation of spouses, which is a function of observed phenotypic correlation, and the underlying mechanism that creates this correlation. To get a sense of what this mechanism may be, we compared tetrachoric correlations across partners of full and half-siblings, and different partners of the same individual.

3.2.1 Results

The twin and adoptee-models were underpowered to estimate heritability of violent crime among women, but using the sibling model identified significant gender differences in the liability to violent crime. Violent crime seemed more heritable among men (A_{male}=59%, A_{female}=28%), whereas the environment was more important among women (C_{male}=13%, C_{female}=23%; E_{male}=28%, E_{female}=49%). The results also indicated that factors loading on C, but not A, varied by sex (ρ_C=0.66, ρ_A=1). This correlation could mean that the environmental factors influencing violent crime are partly different for men and women, but it would also be expected if e.g. some risk factors are of different strength among men and women.

Since we found significant gender differences, and only the sibling model was adequately powered to estimate the heritability of violent crime among women, we proceeded with analyses restricted to men. The pattern of spouse correlations suggested that assortative mating could not be attributed to either complete social homogamy (where mates would be selected on a purely non-genetic factor) or complete primary phenotypic assortment (where mates would be selected based on similarity in violent criminal liability). Under complete social homogamy, the genetic correlation of spouses would be 0. Under complete primary phenotypic assortment, the additive genetic correlation of spouses would be about 0.17. We deemed it likely that the true value was intermediate to these, and Table 2 shows estimates for A and C in the different family models, as a function of spouses' genetic correlation. In the likely range of genetic correlations, the twin and sibling

Table 2. Estimated variance components in different family models, for a range of likely values of parents' genetic correlation.

Model	δ_1	A (95% CI)	C (95% CI)
Twin model	0.05	49% (23%-70%)	16% (0%-38%)
	0.10	52% (25%-70%)	14% (0%-36%)
	0.15	55% (26%-71%)	11% (0%-34%)
Sibling model	0.05	57% (48%-66%)	13% (8%-18%)
	0.10	54% (45%-63%)	13% (8%-17%)
	0.15	51% (43%-60%)	13% (8%-18%)
Adoptee-sibling model	0.05	27% (0%-53%)	4% (0%-20%)
	0.10	25% (0%-50%)	3% (0%-20%)
	0.15	24% (0%-49%)	4% (0%-20%)
Adoptee-parent model	0.05	26% (3%-46%)	0% (0%-18%)
	0.10	25% (3%-44%)	0% (0%-18%)
	0.15	23% (3%-41%)	0% (0%-18%)

Notes: $\delta_1$is the correlation in spouses additive genetic value. For siblings, δ_2, the correlation of consecutive spouses genetic values, is set to 0.4* δ_1, i.e. the value it would take under primary phenotypic assortment

model gives very similar estimates of A and C. The adoptee-models are very similar to each other, but give lower estimates of A and C.

3.2.2 Considerations

This study was the largest study of the heritability of antisocial behavior, one of few to try to adjust for assortative mating, and one of the first to specifically address violent crime. The results could, as all heritability estimates, be criticized for being quite model dependent. As discussed in section 1.3.2, quantitative genetic models rest on untestable assumptions regarding the genetic and environmental correlations of relatives, and unrealistic assumptions of additivity of the effects of individual genes and environmental factors. We attempted to decrease the reliance on these assumptions by comparing models using different types of relatives, and by presenting results for several likely values of genetic assortment. However, the results will still be biased by not being able to estimate variance attributable to deviations from additivity and the covariance of genetic and environmental factors. As always, heritability estimates may serve as a rough guide to the relative importance of genetic and environmental factors, and be informative when comparing e.g. male versus female violent crime.

3.3 PAPER III – INTERPRETATION OF SIBLING COMPARISONS

Paper III is a study of how within-pair estimates from between-within models are influenced by confounders and measurement error. We show, in agreement with previous econometric research, that the sibling comparisons used in co-twin control or sibling difference studies may increase as well as decrease confounding bias, and that attenuation due to random measurement error is stronger for the within-pair estimate. While recognized in economics, we suggest that these

properties of the design have been overlooked in many applications in epidemiology and related disciplines.

For binary exposures, only discordant sibling pairs will contribute to the within-pair estimate. For continuous exposures, all pairs would theoretically contribute to the within-pair estimate but pairs with a large difference in exposure would be more influential, much like a few outliers may have strong influence on a regression coefficient or correlation. This means that the within-pair estimate will be based on pairs selected to be different in exposure, despite the fact that siblings are likely to be similar on the exposure, as they are on most characteristics. This may be viewed as a process where we are selecting (or assigning greater weight to) pairs which differ in causes of the exposure. Some causes of the exposure (those that are strongly familial, i.e. have a high sibling correlation) cannot be very different among siblings, and would not be affected much by this selection. But for causes of the exposure that are less shared, the sibling pairs contributing to the within-pair estimate would be more different from each other than a random pair from the population with the same exposure difference. This includes causes of the exposure that are also causes of the outcome, i.e., confounders. We show the impact this would have under a linear between-within model analytically, and under a logistic between-within model with binary exposure using simulations.

3.3.1 Results

If we let the true causal model be

$$Y_{ij} := \beta_{YX} X_{ij} + \beta_{YC} C_{ij} + \epsilon_{Yij}$$

and

$$X_{ij} := \beta_{XC} C_{ij} + \epsilon_{Xij} ,$$

we have shown that the regression coefficient of regressing X on Y is

$$\beta = \beta_{YX} + \frac{\beta_{YC}\beta_{XC}\sigma_C^2}{\beta_{XC}^2\sigma_C^2 + \sigma_{\epsilon X}^2}$$

and that the within-pair regression coefficient from a between-within model is

$$\beta_W = \beta_{YX} + \frac{\beta_{YC}\beta_{XC}\sigma_C^2}{\beta_{XC}^2\sigma_C^2 + \sigma_{\epsilon X}^2 \frac{1-\rho_{\epsilon X}}{1-\rho_C}} .$$

Both coefficients would be a sum of the true casual effect and a term that is due to confounding. Unlike the unpaired estimate, β, the confounding term of the within-pared estimate, β_W, depends on the sibling correlation in exposure and confounder. When all confounders are perfectly correlated among siblings, the confounding term will vanish and $\beta_W = \beta_{YX}$. When $\rho_C > \rho_{\epsilon X}$ the confounding bias in β_W will be less than in β, and if $\rho_C < \rho_{\epsilon X}$ then β_W is more biased than β. When $\rho_C =$

$\rho_{\epsilon X}$ then $\beta = \beta_W$, even though there may be some confounders perfectly shared by siblings.

In the presence of random measurement error of the exposure, we have shown that the unpaired regression coefficient would be a function of the reliability, γ, of the measurement use to estimate X:

$$\beta^* = \frac{Var(X_{ij})}{Var(X^*_{ij})}\beta = \gamma\beta \,,$$

and the within-pair coefficient, if there is no confounding

$$\beta^*_W(No\ C) = \beta_{YX}\left(1 - \frac{1-\gamma}{1 - Cor(X^*_{i1}, X^*_{i2})}\right),$$

or if there is confounding

$$\beta^*_W = \gamma\left(\beta_{YX}\frac{\beta^2_{XC}\sigma^2_C(1-\rho_C) + \sigma^2_{\epsilon X}(1-\rho_{\epsilon X})}{\beta^2_{XC}\sigma^2_C(1-\gamma\rho_C) + \sigma^2_{\epsilon X}(1-\gamma\rho_{\epsilon X})} + \frac{\beta_{YC}\beta_{XC}\sigma^2_C(1-\rho_C)}{\beta^2_{XC}\sigma^2_C(1-\gamma\rho_C) + \sigma^2_{\epsilon X}(1-\gamma\rho_{\epsilon X})}\right).$$

The ordinary unpaired regression coefficient will be attenuated (biased towards the null) by random measurement error, and the within-pair coefficient even more so, with attenuation increasing as a function of sibling correlation in exposure and outcome.

If the outcome, exposure and confounder are all dichotomous, and the true causal model is a logistic model, we cannot derive exact expressions for the regression coefficients. Instead we simulated this situation, with varying sibling correlations in exposure and outcome, and varying degrees of misclassification of exposure. The results confirmed that the same qualitative conclusions hold for the logistic as for the linear model.

3.3.2 Considerations

Although the paper is written with sibling comparisons such as the co-twin control design in mind, the findings apply equally to others studies using relatives as controls. The formulae and simulation were based on the between-within model. Although we argue that the same conclusions should hold for conditional logistic regression or McNemar tests, this is not illustrated in the paper. However, in all of these models, only the exposure discordant pairs affect the within-pair estimate, so the discussion above still holds.

3.4 PAPER IV – INTELLIGENCE, A CAUSE OF VIOLENT OFFENDING?

In Paper IV, we studied the association of general cognitive ability and violent offending by using a between-within model in full brothers and half-brothers reared together and apart, respectively. We identified brothers and half-brothers

among all men born in Sweden 1961-1975, and who had been enlisted 1980-1993. The outcome was defined as having one or more convictions for a violent offence 1973-2009, and the exposure was the stanine score from the SEB80 intelligence test assessed in early adulthood.

Table 3. Observed and expected probit regression coefficients of general cognitive ability on violent offending, ordinary unpaired analysis and within sibling-pair.

	Probit regression results			Expected within pair under no confounding		
	Unpaired		Within pair			
	Model 1[A]	Model 2[B]	Model 1[A,C]	γ=1	γ=0.9	γ=0.8
Full brothers	-0.19 (-0.19;-0.18)	-0.18 (-0.18;-0.17)	-0.10 (-0.11;-0.09)	-0.19	-0.17	-0.14
Half-brothers reared together	-0.18 (-0.19;-0.17)	-0.17 (-0.19;-0.16)	-0.13 (-0.15;-0.11)	-0.18	-0.17	-0.16
Half-brothers reared apart	-0.18 (-0.19;-0.17)	-0.17 (-0.19;-0.16)	-0.16 (-0.18;-0.14)	-0.18	-0.17	-0.16

Notes: A) Adjusted for birth year. B) Adjusted for birth year and childhood socioeconomic variables: growing up with single mother, family income, and urbanicity. C) Within pair adjustments also included brother's corresponding covariates.

3.4.1 Results

Intelligence was a relatively strong predictor of violent criminal convictions. In the combined sample, 7% of the variance in the liability for violent offending could be attributed to intelligence. In the lowest intelligence stanine category, 20% were convicted of at least one violent offence, in the highest intelligence stanine, only 1%. The association of intelligence and proportion convicted of a violent offence was close to linear on the probit scale. Table 3 summarizes the results.

The within-pair estimates were lower than unpaired estimates for all three sibling groups. In light of the findings in Paper III, we considered that these attenuations may be due to measurement error. The reliability of the SEB80 is likely to be 0.8-0.9. As seen in Table 3, measurement error may explain the attenuation among half-brothers reared apart, but not in the other two groups. We also knew from Paper III that the within estimate may entail both increased and decreased confounding. We argued that confounders were unlikely to create a positive association between intelligence and violent offending, and thus it would be unlikely that the lowered within estimates were due to increased confounding. Based on this, we concluded that the association of intelligence and violent offending was partly confounded by factors strongly shared by brothers reared together. Much of the association remained, however, so this confounding cannot explain the association completely.

3.4.2 Considerations

This is the largest study of the association of general cognitive ability and violent offending to date. Using registered convictions will by necessity combine a

potential influence of intelligence on violent offending with a potential influence of intelligence on being arrested and convicted. This is a prospectively collected data set, and the majority of violent offences were committed after the intelligence measure. However, this does not make us able to make any firm statements on the direction of causation, since both intelligence and antisocial behavior shows substantial within-person stability from childhood and on.

4 DISCUSSION

By capitalizing on a linkage of Swedish national registers, the papers in this thesis are the largest studies yet of the familial aggregation of antisocial behavior (Paper I), of the heritability of antisocial behavior (Paper II), and of the association of general cognitive ability and antisocial behavior (Paper IV). By specifically focusing on violent crime, we targeted a phenotype of great public health and political interest. Before summarizing the findings and implications of the four studies, a few general methodological considerations should be mentioned.

4.1 GENERAL METHODOLOGICAL CONSIDERATIONS

Some methodological considerations have already been addressed in conjunction with each study, but several issues are of more general concern. Papers I, II, and IV were all focused exclusively on registered convictions of violent crime; they may not appropriately have taken truncation and censoring into account; and they relied on paternity information from the MGR. In Papers II-IV, we have used methods which assume that siblings do not have any direct influence on each other, an assumption that may not hold.

4.1.1 Registered crime

In all studies, we relied on registered convictions of violent crime. It is well known that only a fraction of all violent offences will result in a person convicted for the crime. If this fraction was a random sample from the set of all violent offences, it would not have had any real influence on our results other than decreased precision. The estimated heritabilities in Paper II would be the same, and although the associations in Paper I and IV may have been slightly different since both odds ratios and probit regression coefficients may be affected by the outcome's base rate, all comparisons and conclusions would have been unaffected. However, for a crime to result in a conviction, it needs to be reported to the police, the police needs to identify a suspect, who must be charged, and the court must rule against him or her. Based on self-reports, crimes reported to the police are more serious and are less likely to be perpetrated by close relatives [11]. We could easily imagine that there may be discrimination based on e.g. ethnicity or social background affecting which crimes are reported of the police, if not also the outcome of the judicial process. It also seems likely that those who are less able to defend themselves in court, due perhaps to mental health problems or lower education, would be at higher risk of being convicted after committing a crime. While this must be acknowledged, and considered part of the phenotype we have addressed, it should be weighed against the benefit of avoiding reliance on self-reports of crime, and having no issues with response or attrition rates, as inclusion in the register is mandatory.

Using registered violent crime, we were limited in our ability to distinguish crime based on victim type (e.g. a bar fight from domestic violence) or intent (e.g.

instrumental from reactive aggression). This is unfortunate, since we would expect the etiology of different types of violence to be both similar and distinct.

4.1.2 Time-at-risk

Left-truncation due to register start-up and the right censoring due to end of follow-up give different individuals different opportunity to end up in the Register of Criminal Convictions. This was illustrated in Figure 8, which showed that depending on when a person is born, they have widely varying probability to have been registered for a violent crime. This is not only a question of time-at-risk since the rate of convictions for violent crime varies over a person's life time (Figure 1), and different people will be at different age during 1973-2009. Five years at risk would not be the same at age 55-60, as it would be at 20-25. We attempted to correct for this in different ways in Papers I, II and IV.

In Paper I, we performed a nested case-control, matching on birth year and on being at risk when the case was convicted for a crime. This ensured that cases and controls had roughly the same time-at-risk, and experienced this time-at-risk during the same age. In our main results we showed only one estimate for each relative type, but we also showed that familial risks on the OR scale changes with age at first violent crime. For relatives in the same generation, the risks in Fig 1 (in Paper I) could be considered overall estimates of familial risk across different ages at first conviction and socio-economic strata. This may represent an odds increase due to the information in family history alone, with no information on age of sibling, or birth year. Cross-generational relations are more difficult to interpret, since both individuals have to be convicted during a 30-year period. The familial risk of grandparents is based only on combinations of young offenders (the index persons) and older offenders, with an unknown age at their first true conviction, before the start-up of the Crime register. To get an unbiased estimate of grandparental risks, a much longer follow-up time is needed.

In Paper II, we knew that the sibling correlation in birth year would contribute to the observed sibling correlation in violent offending, and tried to compensate for this. First, to make twins and other siblings more comparable in this correlation, we only included siblings born within five years of each other. Second, we adjusted for birth decade as a categorical variable in the analysis. The reason for not actually adjusting for birth year or month was that the analysis time increase linearly with the number of covariate patterns, and compared to the crude estimates, adjusting for birth decade made almost no difference. To further take time-at-risk into account, we excluded immigrants, thereby removing most individuals where birth year was not an acceptable proxy for time-at-risk. However, the adjustment was unnecessarily crude, and we were combining heritability estimates for crimes committed at different stages in life.

In Paper IV, we focused on a well-defined birth cohort, all men born in Sweden 1961-1975. The cohort had reached the age of criminal responsibility when the register started in 1973, and could be followed until they were at least 34 years old, beyond the peak age of violent offending (Figure 1). Birth year was also

included in the analysis as a categorical variable. Although we did not model time-at-risk with e.g. Cox or Poission regression, few individuals were censored before the end of follow up.

In all studies, the outcome was defined as one or more violent offences, and we did not consider the number of crimes. Though there is most certainly information in repeat offending, family-based models for count processes that can properly deal with multiple time scales have not yet been developed.

4.1.3 Paternal discrepancy

Paternity in the MGR is based on a written statement by the mother and father, unless the woman is married, in which case her husband is recorded as the father unless additional information is provided. It is likely that the individual recorded as the father in the MGR is not always the biological father of the child. A review of several international studies of paternal discrepancy, when a man erroneously believes himself to be the biological father of a child, concluded that the rate varied between 0.8%-30%, with a median of about 4% [210]. The higher numbers were, not surprisingly, found in samples where paternity was contested. Paternal discrepancy does not appear to have been studied in Sweden, but we may assume that it is unlikely to be higher than a few percent. This would indicate that some relations we have identified as full-siblings are actually maternal half-siblings, that some paternal half-siblings are not siblings at all, and (probably quite rare) that some half-siblings are actually full siblings. Paternal discrepancy would indicate that several of the familial risks in Paper I would be slightly higher if we had correctly identified biological fathers. However, the true paternity is almost never known in settings where one might consider using family history as a predictor of violent crime, so our estimates are actually closer to what would be practically useful. Paternal discrepancy may lead to underestimation of the heritability in Paper II, but compared to other sources of error, and assumptions encoded in the models, this effect should be very slight. It would not have any real impact in Paper IV, since the interpretation of the within-estimate in different groups is not so sensitive to assumptions regarding genetic correlations.

4.1.4 Sibling interactions

The development of antisocial behavior is often thought to be influenced by associating with antisocial peers (e.g. [4]). It is possible that siblings growing up together would influence each other's liability to perform antisocial acts, or even be partners in crime. We could imagine both a cooperative or imitating interaction with siblings, making them more alike, or a contrasting effect, where the antisocial behavior of one sibling would make the other less likely to engage in antisocial behavior [211]. This would not be a problem for the familial risks in Paper I, where we simply wish to estimate the magnitude of the clustering, regardless of what caused it. However, both the heritability calculation in Paper II, and the interpretation of the sibling comparison in Papers III and IV assume that there is

no effect of the phenotypic value of one sibling on the phenotypic value on the other.

If there is an imitation or cooperation effect, we would expect that siblings who are more similar to begin with, would become even more similar. So the convergence would make monozygotic twins more similar than full siblings, who would in turn be made more similar than half-siblings. Under several theoretical qualifications the presence of sibling convergence would mean that the prevalence of violent crime would be higher among MZ than among DZ twins [211], which we did not observe in Paper II. However, there may be other things separating the groups, such as patterns of non-response (cf Figure 7b). In Paper II, sibling convergence would bias heritability estimates from the twin model upwards, since it would make MZ more alike than DZ. In the adoptee-sibling model, it should be seen as shared environment. In the sibling model it may increase the estimate of shared environment due to the contrasting of paternal and maternal half-siblings, where paternal half-siblings are often not reared together and would thus be affected by only minimal sibling interactions. It may also increase the heritability estimate due to the contrasting of full siblings and maternal half-siblings. The interpretation of sibling comparisons when there is sibling convergence or divergence in exposure or outcome has, to my knowledge, not been explored, and should be considered a research priority.

4.2 FINDINGS AND IMPLICATIONS

4.2.1 Violent crime runs in families

In Paper I, we provided precise estimates of how violent crime runs in families, we argued that the pattern of risk suggested a combination of genetic and environmental factors in explaining this clustering, and showed that there is relatively strong assortative mating for violent offending. Regardless of etiology, the findings in Paper I support that family history may be useful for predicting who might be at risk for committing violent offences. The study was not designed as a prediction model, however, and so cannot tell what the exact predictive value of family history would be compared to other variables. It has previously been reported that family history of CD could improve upon other risk factors, and separate individuals with life-course persistent from adolescence- or childhood-limited antisocial behavior [10]. The results of Paper I suggest that family history may be especially informative for earlier convictions, for female violent crime, and for some specific violent crimes, such as robbery and arson. They also suggest that the etiology of violent crime should probably be seen as a combination of broad processes influencing violent offending over all, and more narrow processes specifically influencing antisocial behavior at different developmental stages, affecting men and women differently, and influencing propensity to specific types of violent crime [212]. As reviewed in the introduction, this would be in agreement with previous findings for other less severe antisocial behaviors.

In Paper II, we showed the familial clustering of violent crime can be explained by a combination of genetic and environmental influences on the liability to commit violent offences. We found that sibling and twin models produced estimates of the heritability of being convicted for violent crime that did not differ much from general antisocial behavior, while adoptee models gave lower estimates of both the heritability and the family shared environment. This difference may be due to the rarity of violent crime among adopted parents, who are carefully screened before being allowed to adopt a child [190]. However, this would go against one previous study which reported that such *range restriction*, though quite strong, did not influence sibling correlations in that sample [213]. We also showed that the assortative mating for violent offending could not be attributed to either social homogamy or primary phenotypic assortment. If the mate selection were due to some combination of these processes, we showed that it would only have a small effect on estimates from the adoptee-model, no effect on the environmental estimates from the sibling model, and modest influence on heritability estimates from sibling and twin models, and environmental estimates from the twin model. Although adjustment for assortative mating did not change any of the conclusions in this study, they may be of consequence in more elaborate quantitative genetic designs, and should not be overlooked.

Finding substantial heritability of the liability to violent offending implies that inherited characteristics should be considered as possible causes of violent offending, and confounders of other established risk factors. However, it does not necessarily imply that specific genes are likely to improve our ability to predict future violence. It is a general finding for complex traits that individual alleles have very weak influence on the trait; most often, no alleles are found that account for even 1 % of the phenotypic variance. It has been argued that this makes it unlikely, even in the context of human height, that genotyping will, in the near future, aid us in making prediction models that outperform established risk factors, and the simple "Victorian" practice of predicting a child's phenotype by averaging his parents' phenotypic values [214].

In terms of better understanding the etiology of crime, however, genetics provides a fascinating perspective. Though few alleles have been identified in studies of conduct disorder or delinquency, it is not true to say that we know nothing of the genetics of antisocial behavior. In a recent meta-analysis of GWA studies on brain volume, an allele (rs10784502) previously found to be associated with height was significantly associated to intracranial volume [215]. Since intracranial brain volume is known to be positively correlated to intelligence, the researchers tested the allele against full-scale IQ in a subsample, and found a significant association, driven by an association with performance IQ. The allele is situated in an intron in the gene HMGA2, which encodes a chromatin-associated protein involved in cell growth [216]. If this is indeed a gene which influences IQ through its effect on brain volume; then if IQ has an influence on individual risk for violent crime, it is

also a gene with an influence on violent crime. Along the same lines, if we are persuaded by arguments that schizophrenia and bipolar disorder may have a causal effect on the propensity for violent offending, the genes identified for these disorders are also likely to be genes for violent crime [217, 218].

In better understanding the proximal causes of violence, and better understanding the causes of these causes, we will get a fuller picture of what leads to violent crime, including the distal genetic factors. But this still leaves us with the problem of showing that a particular risk factor is causally related to violent offending.

4.2.2 Sibling control studies may be difficult to interpret

Studies based on using relatives as controls has been promoted as helpful for addressing issues of causality, able to separate a causal effect from confounding due to shared environment and genetic factors [6, 79, 151]. However, the models have received little attention from statistical or methodological standpoints, and there has been confusion on how they should be implemented and interpreted.

In Paper III, we showed that within-pair estimates from sibling comparisons may in some situations be more biased than ordinary unpaired estimates from traditional observational study designs. In particular, we showed that if there is measurement error, the within-pair estimate is expected to be lower (closer to the null) than the ordinary estimate, even if there is no confounding. If there is confounding, the within-pair estimate may be either less or more confounded than the unpaired estimate. When siblings are less similar (less correlated) in exposure than in confounders, confounding bias will be lower within-pair. When siblings are more similar in exposure than in confounders, confounding bias will be higher within-pair. When siblings are similarly correlated in exposure and confounders, the within-pair estimate will be close to the unpaired estimate even though confounders are to some extent shared by siblings. This makes the interpretation of sibling comparisons more complicated than often acknowledged.

Based on Paper III, it seems we might need to interpret studies using relatives as controls in light of the observed pair correlation in exposure; reasonable values for the pair correlation of confounders; knowledge on whether confounding is likely to create a positive or negative association; and how well the observed exposure measures the causal exposure. To be able to decide what alternative explanations may explain an observed within-pair estimate; sibling control studies may need considerable statistical power. This seems particularly important for being able to separate attenuation in the within-pair estimate due to measurement error from a reduction in the association due to reduced (or increased) confounding bias.

Due to the difficulty in interpreting within-pair estimates in the presence of confounders less than perfectly shared by the pair, sibling comparisons may be

most useful in situations where there is a believable hypothesis that factors shared perfectly by the pair may explain the observed association. This would, for instance, include situations where socioeconomic background, ethnicity, or rearing environment has been suggested as strong confounders of a putatively causal association. Unless the association is completely removed within pairs perfectly correlated on the suspected confounder, this confounder cannot explain the association completely. It seems more difficult to use sibling controls to get an unbiased estimate of the effect free from such factors when they are only partially responsible for the association. It also seems difficult to adjust for so called genetic confounding. Although MZ twins are genetically identical, they are not identical on even very heritable traits. Even when an association remains unchanged within MZ twin pairs, this association may well be completely confounded by highly heritable factors, if they are similarly correlated as the exposure.

4.2.3 Intelligence and violent crime

In Paper IV, we showed that general cognitive ability was inversely associated with having been convicted of one or more violent crimes. Despite focusing on convictions for violent crime, the strength of the association was close to that reported previously for self-reported antisocial behavior. By comparing this association with the within-pair association among full brothers, half-brothers registered as reared together, and half-brothers reared apart, we hoped to gain some insight into the nature of the intelligence- violent crime association.

The within-pair associations were all lower than the unpaired estimates. While measurement error could explain the attenuation among half-brothers reared apart, the attenuations were stronger than expected among half-brothers reared together and full brothers, indicating increased or decreased confounding bias. Arguing that overall confounding was unlikely to create an association between high intelligence and violent crime, we concluded that the association was partly confounded by factors shared by brothers reared together. We could speculate that these confounders are connected to the early shared childhood factors, possible related to parenting practices or even neglect/abuse, but unfortunately we lacked additional information needed to test these hypotheses. Much of the association remained within full-brother pairs, however, so these shared confounding factors could not explain the association completely.

It is important to note that the sibling comparison is not able to separate the effect if intelligence from other factors strongly correlated to it. For instance, it has been argued that the association may be due to other psychological traits, such as executive functioning or impulsivity [145]. Since these factors are likely to have a similar pattern of correlation over the half- and full-brother relations, it is possible that they may better explain the intelligence-crime association.

4.3 CONCLUSIONS

In this thesis, I have shown that violent crime runs in families, and argued that this is due to both genetic and environmental factors influencing the propensity to violent offending. The familial aggregation of violent crime and the etiological importance of both genetic factors and factors in the rearing environment should be acknowledged in criminological research.

The complications pointed out in Paper III should be considered when interpreting results from studies using relatives as controls, and should perhaps temper claims of these designs' ability to address issues of causality.

In light of Paper III, I could in Paper IV conclude that most of the intelligence-offending association is not due to confounding by childhood environment, but I was not able to make any statement on the possible confounding by other similarly heritable factors, or so called genetic confounding. If general cognitive ability does indeed have a causal impact on violent crime, this would explain part of the heritability of violent crime. As factors, including genes, influencing intelligence are identified; we would expect them to also be causes of violent crime. Causation aside, it should be recognized by policy makers and managers in the judicial system that individuals convicted of violent offending have weaker cognitive resources than the general population.

5 ACKNOWLEDGEMENTS

This thesis was supported by the Swedish Prison and Probation Service, by employing me half-time to work on it. Thus, my past four years have been split between time as a PhD-student at the Department of Medical Epidemiology and Biostatistics at Karolinska Institutet, and time as a project manager at the Swedish Prison and Probation Service R&D. In both organizations, I have had the pleasure of working with brilliant, dedicated, and genuinely interesting colleagues, many of whom I consider dear friends. In particular, I would like to acknowledge the following people.

Niklas Långström, my main supervisor, and as head of the Prison and Probation Service R&D Stockholm Unit, also my boss. Our collaboration may have started with non-heterosexual sex, but it quickly turned to violence. Thank you for believing in me from the start, for encouraging me even when our interests did not perfectly align, and for teaching me how to find the story in an analysis. And thank you for giving me the opportunity to pursue my PhD while employed by the Prison and Probation Service. The experience taught me so much; scientifically and personally!

Yudi Pawitan, my co-supervisor. Thank you for making me pick up a pencil and start doing math again. I may not have understood all of it at the time, or even now, but your course on likelihood-based inference forced me to go from using programs to writing my own.

Paul Lichtenstein, my co-supervisor, and head of the Swedish Twin Registry. Thank you for creating such a vibrant research environment, over-flowing with rich and unique data sets. And thank you for letting me play with them! The projects in this thesis would not have been possible without you.

All the office roommates I've had over the years; **Christina Persson, Petra Buxrud, Yasmina Molero-Samuelson, Lovisa Högberg, Linda Halldner Henriksson**, and, above all, **Therese Ljung**. Thank you for putting up with me, my dying plants, and my music!

The Epi-Gang: **Christin Bexelius**, **Arvid Sjölander**, **Anna Svensson**, and **Sara Öberg**. You have been my mentors, teachers, and friends. Thank you for proving that brilliant scientists can also be brilliant dinner company. And thank you for teaching me epidemiology!

The iron core of PubMEB: **Caroline Dietrich**, **Sandra Eloranta**, and **Robert Karlsson**. I know you've gotten capable beer drinkers to replace me, but should you ever need help finishing the fridge, remember I'll be just across campus. Beer! Pizza! Music!

My past and present colleagues at Kriminalvården FOU Stockholm: **Petra Buxrud**, **Axel Fors**, **Mats Forsman**, **Ingrid Freij**, **Ralf Kuja-Halkola**, **Bengt Holmgren**, **Lena Lundholm**, **Andrea Mogues**, **Elizabeth Nordén**, **Jenny Yourstone**, and **Diana Wenander**. Thank you for, in seminars, coffee breaks, and the occasional AW, lending me your eyes to look at criminology, academia, research, and life. Your perspectives have made me wiser.

The MEB PhD-student gang. **Stephanie Bonn**, **Sara Christensen**, **Miriam Elfström**, **Adina Feldman**, **Martin Fransson**, **Alexander Grankvist**, **Lovisa Högberg**, **Therese Ljung**, **Lisa Möller**, **Christina Persson**, **Iffat Rahman**, **Maria Sandberg**, **Karin Sundström**, and others. Thank you for being my friends inside MEB, and out. Y'all don't be strangers, now!

My parents, **Johnny** and **Ulla**, and my brothers **Björn** and **Peter** with families. Thank you for keeping me grounded, and always providing a fresh point of view.

Finally, **Fredrik Havasi**. Thank you for your economist's perspective, for tending the base camp, and for sticking around. I love you, babe!

6 REFERENCES

1. Krug, EG, et al., *The world report on violence and health.* Lancet, 2002. **360**(9339): p. 1083-88.
2. Sethi, D, et al., *European report on preventing violence and knife crime among young people*, 2010, The Regional Office for Europe of the World Health Organization.
3. Stattin, H and Magnusson, D, *Stability and Change in Criminal Behavior up to Age 30.* Brit J Criminol, 1991. **31**(4): p. 327-46.
4. Farrington, DP, *A criminological research agenda for the next millennium.* Int J Offender Ther, 1999. **43**(2): p. 154-67.
5. Cohen, M and Piquero, A, *New Evidence on the Monetary Value of Saving a High Risk Youth.* J Quant Criminol, 2009. **25**(1): p. 25-49.
6. Moffitt, TE, *The new look of behavioral genetics in developmental psychopathology: gene-environment interplay in antisocial behaviors.* Psychol Bull, 2005. **131**(4): p. 533-54.
7. McGue, M, Osler, M, and Christensen, K, *Causal Inference and Observational Research: The Utility of Twins.* Perspect Psychol Sci, 2010. **5**(5): p. 546-56.
8. Widom, CS, *Does violence beget violence? A critical examination of the literature.* Psychol Bull, 1989. **106**(1): p. 3-28.
9. Farrington, DP, et al., *The concentration of offenders in families, and family criminality in the prediction of boys' delinquency.* J Adolesc, 2001. **24**(5): p. 579-96.
10. Odgers, CL, et al., *Predicting prognosis for the conduct-problem boy: can family history help?*J Am Acad Child Adolesc Psychiatry, 2007. **46**(10): p. 1240-9.
11. Swedish National Council for Crime Prevention, *Brottsoffers benägenhet att anmäla brott, rapport 2008:12*, 2008, Brottsförebyggande rådet: Stockholm.
12. Westfelt, L, *Svensk brottslighet i internationell belysning*, in *Brottsutvecklingen i Sverige fram till 2007.*, L. Ekström, Editor 2008, Swedish National Council for Crime Prevention: Stockholm. p. 448-73.
13. Swedish National Council for Crime Prevention, *Kriminalstatistik 2010*, 2011, Brottsförebyggande rådet: Stockholm.
14. Federal Bureau of Investigation. *Crime in the United States - Table 1*. [cited 2012 April 16th]; Available from: http://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2010/crime-in-the-u.s.-2010/tables/10tbl01.xls.
15. Statistics Sweden, *Beskriving av befolkningen i Sverige 2008* [*Description of the Population in Sweden 2008]*, 2009. Available from: http://www.scb.se/statistik/_publikationer/BE0101_2008A01_BR_BE0109TEXT.pdf
16. Martens, P and Holmberg, S, *Brottslighet bland personer födda i Sverige och i utlandet [Crime among individuals born in Sweden and abroad]*, 2005, The Swedish National Council for Crime Prevention: Stockholm.
17. Beckley, A, *Correlates of War? Toward an understanding of nativity-based variation in immigrant offending*, 2012, Working paper, Stockholm university: Stockholm.
18. Swedish National Council for Crime Prevention, *Brottsutvecklingen i Sverige fram till år 2007*, 2008, Brottsförebyggande rådet: Stockholm.
19. Von Hofer, H, *Brott och straff i Sverige*, in *Kriminologiska institutionens rapportserie*, F. Estrada, Editor 2011, Kriminologiska Institutionen: Stockholm.
20. Rhee, SH and Waldman, ID, *Genetic and environmental influences on antisocial behavior: A meta-analysis of twin and adoption studies.* Psychol Bull, 2002. **128**(3): p. 490-529.
21. APA, *Diagnostic and statistical manual of mental disorders 4th edn text revision*2000, Washington, DC: American Psychiatric Association.
22. Hare, RD, et al., *Psychopathy and the predictive validity of the PCL-R: an international perspective.* Behav Sci Law, 2000. **18**(5): p. 623-45.
23. Healy, W, and Bronner, AF, *New light on delinquency and its treatment: results of a research conducted for the Institute of human relations*1936: Pub. for the Institute of human relations by Yale university press.
24. Burt, SA and Neiderhiser, JM, *Aggressive Versus Nonaggressive Antisocial Behavior: Distinctive Etiological Moderation by Age.* Developmental Psychology, 2009. **45**(4): p. 1164-76.
25. Baker, LA, et al., *Differential Genetic and Environmental Influences on Reactive and Proactive Aggression in Children.* J Abnorm Child Psych, 2008. **36**(8): p. 1265-78.

26. Tuvblad, C, et al., *Genetic and Environmental Stability Differs in Reactive and Proactive Aggression.* Aggressive Behav, 2009. **35**(6): p. 437-52.
27. Burt, SA and Klump, KL, *The Etiological Moderation of Aggressive and Nonaggressive Antisocial Behavior by Age*. Twin Res Hum Genetics, 2009. **12**(4): p. 343-50.
28. Anwar, S, et al., *Is arson the crime most strongly associated with psychosis?--A national case-control study of arson risk in schizophrenia and other psychoses.* Schizophr Bull, 2011. **37**(3): p. 580-6.
29. Fergusson, DM, Horwood, LJ, and Nagin, DS, *Offending trajectories in a New Zealand birth cohort.* Criminology, 2000. **38**(2): p. 525-51.
30. Soothill, K, et al., *Changing patterns of offending behaviour among young adults.* Brit J Criminol, 2008. **48**(1): p. 75-95.
31. Lussier, P, *The criminal activity of sexual offenders in adulthood: revisiting the specialization debate.* Sexual Abuse, 2005. **17**(3): p. 269-92.
32. Wanklyn, SG, et al., *Can We Distinguish Juvenile Violent Sex Offenders, Violent Non-Sex Offenders, and Versatile Violent Sex Offenders Based on Childhood Risk Factors?* J Interpers Violence, 2012.
33. Kjellgren, C, et al., *Sexually coercive behavior in male youth: population survey of general and specific risk factors.* Arch Sex Behav, 2010. **39**(5): p. 1161-9.
34. Krueger, RF, et al., *Etiologic connections among substance dependence, antisocial behavior, and personality: Modeling the externalizing spectrum.* J Abnorm Psychol, 2002. **111**(3): p. 411-24.
35. Burt, SA, *Are there meaningful etiological differences within antisocial behavior? Results of a meta-analysis.* Clinical Psychology Review, 2009. **29**(2): p. 163-78.
36. Hicks, BM, et al., *Family transmission and heritability of externalizing disorders: a twin-family study.* Arch Gen Psychiatry, 2004. **61**(9): p. 922-8.
37. Molero Samuelson, Y, *Antisocial behaviour over the life course among females and males treated for substance misuse*. Doctoral thesis, 2011, Karolinska Institutet: Stockholm.
38. Brame, B, Nagin, DS, and Tremblay, RE, *Developmental trajectories of physical aggression from school entry to late adolescence.* Journal of Child Psychology and Psychiatry and Allied Disciplines, 2001. **42**(4): p. 503-12.
39. Fergusson, DM and Horwood, LJ, *Early conduct problems and later life opportunities.* J Child Psychol Psychiatry, 1998. **39**(8): p. 1097-108.
40. Fergusson, DM, Boden, JM, and Horwood, LJ, *Situational and generalised conduct problems and later life outcomes: evidence from a New Zealand birth cohort.* J Child Psychol Psychiatry, 2009. **50**(9): p. 1084-92.
41. Ehrensaft, MK, et al., *Intergenerational transmission of partner violence: a 20-year prospective study.* J Consult Clin Psychol, 2003. **71**(4): p. 741-53.
42. Simonoff, E, et al., *Predictors of antisocial personality. Continuities from childhood to adult life.* The British journal of psychiatry : the journal of mental science, 2004. **184**: p. 118-27.
43. Gendreau, P, Little, T, and Goggin, C, *A meta-analysis of the predictors of adult offender recidivism: What works.* Criminology, 1996. **34**(4): p. 575-607.
44. Moffitt, TE, *Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy.* Psychol Rev, 1993. **100**(4): p. 674-701.
45. Boden, JM, Fergusson, DM, and Horwood, LJ, *Risk factors for conduct disorder and oppositional/defiant disorder: evidence from a New Zealand birth cohort.* J Am Acad Child Adolesc Psychiatry, 2010. **49**(11): p. 1125-33.
46. Ge, XJ, Donnellan, MB, and Wenk, E, *The development of persistent criminal offending in males.* Crim Justice Behav, 2001. **28**(6): p. 731-55.
47. Raine, A, et al., *Neurocognitive impairments in boys on the life-course persistent antisocial path.* J Abnorm Psychol, 2005. **114**(1): p. 38-49.
48. Donnellan, MB, Ge, XJ, and Wenk, E, *Cognitive abilities in adolescent-limited and life-course-persistent criminal offenders.* J Abnorm Psychol, 2000. **109**(3): p. 396-402.
49. Odgers, CL, et al., *Female and male antisocial trajectories: from childhood origins to adult outcomes.* Dev Psychopathol, 2008. **20**(2): p. 673-716.
50. Tremblay, RE, et al., *Physical aggression during early childhood: Trajectories and predictors.* Pediatrics, 2004. **114**(1): p. E43-E50.
51. Nagin, DS, Farrington, DP, and Moffitt, TE, *Life-Course Trajectories of Different Types of Offenders.* Criminology, 1995. **33**(1): p. 111-39.
52. Farrington, DP, Ttofi, MM, and Coid, JW, *Development of Adolescence-Limited, Late-Onset, and Persistent Offenders From Age 8 to Age 48.* Aggressive Behav, 2009. **35**(2): p. 150-63.

53. Nagin, DS and Tremblay, RE, *What has been learned from group-based trajectory modeling? Examples from physical aggression and other problem behaviors.* Ann Am Acad Polit Ss, 2005. **602**: p. 82-117.
54. Cote, SM, et al., *The development of physical aggression from toddlerhood to pre-adolescence: a nation wide longitudinal study of Canadian children.* J Abnorm Child Psych, 2006. **34**(1): p. 71-85.
55. Besemer, S and Farrington, DP, *Intergenerational transmission of criminal behaviour: Conviction trajectories of fathers and their children.* Eur J Criminol, 2012. **9**(2): p. 120-41.
56. Thornberry, TP, *Explaining multiple patterns of offending across the life and course and across generations.* Ann Am Acad Polit Ss, 2005. **602**: p. 156-95.
57. Sampson, RJ and Laub, JH, *A life-course view of the development of crime.* Ann Am Acad Polit Ss, 2005. **602**: p. 12-45.
58. Brame, R, Bushway, SD, and Paternoster, R, *Examining the prevalence of criminal desistance.* Criminology, 2003. **41**(2): p. 423-48.
59. Hay, DF, *The gradual emergence of sex differences in aggression: alternative hypotheses.* Psychol Med, 2007. **37**(11): p. 1527-37.
60. Storvoll, EE and Wichstrom, L, *Gender differences in changes in and stability of conduct problems from early adolescence to early adulthood.* J Adolesc, 2003. **26**(4): p. 413-29.
61. Odgers, C and Moretti, MM, *Aggressive and Antisocial Girls: Research Update and Challenges.* Int J Forens Ment Health, 2002. **1**(2): p. 103-19.
62. East, K and Campbell, S, *Aspects of crime: Young offenders 1999*, H.O. UK, Editor 1999.
63. Javdani, S, Sadeh, N, and Verona, E, *Expanding our lens: female pathways to antisocial behavior in adolescence and adulthood.* Clinical psychology review, 2011. **31**(8): p. 1324-48.
64. Cloninger, CR, et al., *Implications of Sex-Differences in Prevalences of Antisocial Personality, Alcoholism, and Criminality for Familial Transmission.* Arch Gen Psychiatry, 1978. **35**(8): p. 941-51.
65. Moretti, MM, Catchpole, RE, and Odgers, C, *The dark side of girlhood: recent trends, risk factors and trajectories to aggression and violence.* Can Child Adolesc Psychiatr Rev, 2005. **14**(1): p. 21-5.
66. Krakowski, M and Czobor, P, *Gender differences in violent behaviors: Relationship to clinical symptoms and psychosocial factors.* Am J Psychiat, 2004. **161**(3): p. 459-65.
67. Burt, SA, et al., *Timing of menarche and the origins of conduct disorder.* Arch Gen Psychiatry, 2006. **63**(8): p. 890-6.
68. Putkonen, A, et al., *The quantitative risk of violent crime and criminal offending: a case-control study among the offspring of recidivistic Finnish homicide offenders.* Acta Psychiatrica Scandinavica, Supplementum, 2002(412): p. 54-7.
69. Bijleveld, CC and Wijkman, M, *Intergenerational continuity in convictions: a five-generation study.* Crim Behav Ment Health, 2009. **19**(2): p. 142-55.
70. Putkonen, A, et al., *Transmission of violent offending and crime across three generations.* Soc Psychiatry Psychiatr Epidemiol, 2007. **42**(2): p. 94-9.
71. van de Rakt, M, Nieuwbeerta, P, and de Graaf, ND, *Like Father, Like Son: The Relationships between Conviction Trajectories of Fathers and their Sons and Daughters.* Br J Criminol, 2008. **48**(4): p. 538-56.
72. Pears, KC and Capaldi, DM, *Intergenerational transmission of abuse: a two-generational prospective study of an at-risk sample.* Child Abuse Negl, 2001. **25**(11): p. 1439-61.
73. Dubow, EF, Huesmann, LR, and Boxer, P, *Theoretical and methodological considerations in cross-generational research on parenting and child aggressive behavior.* J Abnorm Child Psychol, 2003. **31**(2): p. 185-92; discussion 201-3.
74. Serbin, LA and Karp, J, *The intergenerational transfer of psychosocial risk: mediators of vulnerability and resilience.* Annu Rev Psychol, 2004. **55**: p. 333-63.
75. Clark, DB, et al., *Psychopathology risk transmission in children of parents with substance use disorders.* Am J Psychiatry, 2004. **161**(4): p. 685-91.
76. Jonsson, G, *Delinquent boys, their parents and grandparents.* Acta Psychiatrica Scandinavica, Supplementum, 1967. **43(SUPPL. 195)**.
77. Conger, RD, et al., *Angry and aggressive behavior across three generations: A prospective, longitudinal study of parents and children.* JAbnorm Child Psych, 2003. **31**(2): p. 143-60.
78. Delsol, C and Margolin, G, *The role of family-of-origin violence in men's marital violence perpetration.* Clin Psychol Rev, 2004. **24**(1): p. 99-122.

79. Jaffee, SR, Strait, LB, and Odgers, CL, *From correlates to causes: can quasi-experimental studies and statistical innovations bring us closer to identifying the causes of antisocial behavior?* Psychol Bull, 2012. **138**(2): p. 272-95.
80. Junginger, J, *Psychosis and violence: the case for a content analysis of psychotic experience.* Schizophr Bull, 1996. **22**(1): p. 91-103.
81. Anckarsater, H, et al., *Mental disorder is a cause of crime: The cornerstone of forensic psychiatry.* Int J Law Psych, 2009. **32**(6): p. 342-47.
82. Mullen, PE, *Schizophrenia and violence: from correlations to preventive strategies.* Advances in Psychiatric Treatment, 2006. **12**: p. 239-48.
83. Mullen, PE, *Mental Disorder and Dangerousness.* Aust Nz J Psychiat, 1984. **18**(1): p. 8-17.
84. Monahan, J and Steadman, HJ, *Crime and Mental Disorder: An Epidemiological Approach.* Crime and Justice, 1983. **4**: p. 145-89.
85. Brennan, PA, Mednick, SA, and Hodgins, S, *Major mental disorders and criminal violence in a Danish birth cohort.* Arch Gen Psychiatry, 2000. **57**(5): p. 494-500.
86. Fazel, S, et al., *Schizophrenia, substance abuse, and violent crime.* Jama, 2009. **301**(19): p. 2016-23.
87. Fazel, S, et al., *Bipolar disorder and violent crime: time at risk reanalysis.* Arch Gen Psychiatry, 2010. **67**(12): p. 1325-6.
88. Fazel, S, et al., *Bipolar disorder and violent crime: new evidence from population-based longitudinal studies and systematic review.* Arch Gen Psychiatry, 2010. **67**(9): p. 931-8.
89. Steadman, HJ, et al., *Violence by people discharged from acute psychiatric inpatient facilities and by others in the same neighborhoods.* Arch Gen Psychiatry, 1998. **55**(5): p. 393-401.
90. Swanson, JW, et al., *Violence and psychiatric disorder in the community: evidence from the Epidemiologic Catchment Area surveys.* Hosp Community Psychiatry, 1990. **41**(7): p. 761-70.
91. Tiihonen, J, et al., *Specific major mental disorders and criminality: a 26-year prospective study of the 1966 northern Finland birth cohort.* Am J Psychiatry, 1997. **154**(6): p. 840-5.
92. Torrey, EF, et al., *The MacArthur Violence Risk Assessment Study revisited: two views ten years after its initial publication.* Psychiatr Serv, 2008. **59**(2): p. 147-52.
93. Wallace, C, Mullen, PE, and Burgess, P, *Criminal offending in schizophrenia over a 25-year period marked by deinstitutionalization and increasing prevalence of comorbid substance use disorders.* Am J Psychiatry, 2004. **161**(4): p. 716-27.
94. Walsh, E, Buchanan, A, and Fahy, T, *Violence and schizophrenia: examining the evidence.* Br J Psychiatry, 2002. **180**: p. 490-5.
95. Fazel, S, et al., *Schizophrenia and violence: systematic review and meta-analysis.* PLoS Med, 2009. **6**(8): p. e1000120.
96. Fazel, S and Grann, M, *The population impact of severe mental illness on violent crime.* Am J Psychiatry, 2006. **163**(8): p. 1397-403.
97. Hoaken, PNS and Stewart, SH, *Drugs of abuse and the elicitation of human aggressive behavior.* Addict Behav, 2003. **28**(9): p. 1533-54.
98. Boles, SM and Miotto, K, *Substance abuse and violence - A review of the literature.* Aggress Violent Beh, 2003. **8**(2): p. 155-74.
99. Grann, M and Fazel, S, *Substance misuse and violent crime: Swedish population study.* Bmj, 2004. **328**(7450): p. 1233-4.
100. Green, KM, et al., *Does heavy adolescent marijuana use lead to criminal involvement in adulthood? Evidence from a multiwave longitudinal study of urban African Americans.* Drug Alcohol Depen, 2010. **112**(1-2): p. 117-25.
101. Green, KM, et al., *Association Between Adolescent Drinking and Adult Violence: Evidence From a Longitudinal Study of Urban African Americans.* J Stud Alcohol Drugs, 2011. **72**(5): p. 701-10.
102. Haggård-Grann, U, et al., *The role of alcohol and drugs in triggering criminal violence: a case-crossover study.* Addiction, 2006. **101**(1): p. 100-08.
103. Parker, RN and Auerhahn, K, *Alcohol, drugs, and violence.* Annu Rev Sociol, 1998. **24**: p. 291-311.
104. Doherty, EE, Green, KM, and Ensminger, ME, *Investigating the long-term influence of adolescent delinquency on drug use initiation.* Drug Alcohol Depen, 2008. **93**(1-2): p. 72-84.
105. Disney, ER, et al., *Effects of ADHD, conduct disorder, and gender on substance use and abuse in adolescence.* Am J Psychiat, 1999. **156**(10): p. 1515-21.

106. McGue, M and Iacono, WG, *The association of early adolescent problem behavior with adult psychopathology.* Am J Psychiat, 2005. **162**(6): p. 1118-24.
107. Fergusson, DM, Horwood, LJ, and Ridder, EM, *Conduct and attentional problems in childhood and adolescence and later substance use, abuse and dependence: Results of a 25-year longitudinal study.* Drug Alcohol Depen, 2007. **88**: p. S14-S26.
108. Kim-Cohen, J, et al., *Prior juvenile diagnoses in adults with mental disorder - Developmental follow-back of a prospective-longitudinal cohort.* Arch Gen Psychiatry, 2003. **60**(7): p. 709-17.
109. White, HR, et al., *Developmental associations between substance use and violence.* Dev Psychopathol, 1999. **11**(4): p. 785-803.
110. Fergusson, DM, John Horwood, L, and Ridder, EM, *Show me the child at seven II: childhood intelligence and later outcomes in adolescence and young adulthood.* J Child Psychol Psyc, 2005. **46**(8): p. 850-58.
111. Neisser, U, et al., *Intelligence: Knowns and unknowns.* Am Psychol, 1996. **51**(2): p. 77-101.
112. Gottfredson, LS, *Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography (Reprinted from The Wall Street Journal, 1994).* Intelligence, 1997. **24**(1): p. 13-23.
113. Carroll, JB, *Psychometrics, intelligence, and public perception.* Intelligence, 1997. **24**(1): p. 25-52.
114. Carlstedt, B, *Cognitive abilities - aspects of structure, process and measurement*. Doctoral thesis,. 2000, Göteborg: Acta Universitatis Gothoburgensis.
115. McGrew, KS, *CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research.* Intelligence, 2009. **37**(1): p. 1-10.
116. Horn, JL and Cattell, RB, *Refinement and test of the theory of fluid and crystallized general intelligences.* J Educ Psychol, 1966. **57**(5): p. 253-70.
117. Rushton, JP and Ankney, CD, *Whole brain size and general mental ability: a review.* Int J Neurosci, 2009. **119**(5): p. 691-731.
118. Luders, E, et al., *Neuroanatomical Correlates of Intelligence.* Intelligence, 2009. **37**(2): p. 156-63.
119. Deary, IJ, Penke, L, and Johnson, W, *The neuroscience of human intelligence differences.* Nat Rev Neurosci, 2010. **11**(3): p. 201-11.
120. Haworth, CM, et al., *A twin study of the genetics of high cognitive ability selected from 11,000 twin pairs in six studies from four countries.* Behav Genet, 2009. **39**(4): p. 359-70.
121. Haworth, CM, et al., *The heritability of general cognitive ability increases linearly from childhood to young adulthood.* Mol Psychiatr, 2010. **15**(11): p. 1112-20.
122. Hirschi, T and Hindelang, MJ, *Intelligence and Delinquency - Revisionist Review.* Am Sociol Rev, 1977. **42**(4): p. 571-87.
123. Simons, RL, *Meaning of Iq Delinquency Relationship.* Am Sociol Rev, 1978. **43**(2): p. 268-70.
124. Menard, S and Morse, BJ, *A Structuralist Critique of the Iq-Delinquency Hypothesis - Theory and Evidence.* Am J Sociol, 1984. **89**(6): p. 1347-78.
125. Harry, J and Minor, WW, *Intelligence and Delinquency Reconsidered - a Comment.* Am J Sociol, 1986. **91**(4): p. 956-62.
126. Moffitt, TE and Silva, PA, *Iq and Delinquency - a Direct Test of the Differential Detection Hypothesis.* J Abnorm Psychol, 1988. **97**(3): p. 330-33.
127. Farrington, DP, *Early predictors of adolescent aggression and adult violence.* Violence Vict, 1989. **4**(2): p. 79-100.
128. Barker, ED, et al., *Developmental trajectories of male physical violence and theft - Relations to neurocognitive performance.* Arch Gen Psychiatry, 2007. **64**(5): p. 592-99.
129. Kennedy, TD, Burnett, KF, and Edmonds, WA, *Intellectual, behavioral, and personality correlates of violent vs. non-violent juvenile offenders.* Aggress Behav, 2011. **37**(4): p. 315-25.
130. Guay, JP, Ouimet, M, and Proulx, J, *On intelligence and crime: A comparison of incarcerated sex offenders and serious non-sexual violent criminals.* Int J Law Psychiat, 2005. **28**(4): p. 405-17.
131. Hodgins, S, *Mental Disorder, Intellectual Deficiency, and Crime - Evidence from a Birth Cohort.* Arch Gen Psychiatry, 1992. **49**(6): p. 476-83.
132. McGloin, JM, Pratt, TC, and Maahs, J, *Rethinking the IQ-delinquency relationship: A longitudinal analysis of multiple theoretical models.* Justice Q, 2004. **21**(3): p. 603-35.

133. Koenen, KC, et al., *Genetic influences on the overlap between low IQ and antisocial behavior in young children.* J Abnorm Psychol, 2006. **115**(4): p. 787-97.
134. Isen, J, *A meta-analytic assessment of Wechsler's P > V sign in antisocial populations.* Clinical Psychology Review, 2010. **30**(4): p. 423-35.
135. Moffitt, TE, et al., *Socioeconomic-Status, Iq, and Delinquency.* J Abnorm Psychol, 1981. **90**(2): p. 152-56.
136. Lynam, D, Moffitt, T, and Stouthamer-Loeber, M, *Explaining the relation between IQ and delinquency: class, race, test motivation, school failure, or self-control?* J Abnorm Psychol, 1993. **102**(2): p. 187-96.
137. Levine, SZ, *Elaboration on the association between IQ and parental SES with subsequent crime.* Pers Indiv Differ, 2011. **50**(8): p. 1233-37.
138. Schonfeld, IS, et al., *Conduct Disorder and Cognitive-Functioning - Testing 3 Causal Hypotheses.* Child Dev, 1988. **59**(4): p. 993-1007.
139. Mõttus, R, et al., *Longitudinal associations of cognitive ability, personality traits and school grades with antisocial behaviour.* European Journal of Personality, 2011: p. n/a-n/a.
140. Ward, DA and Tittle, CR, *Iq and Delinquency - a Test of 2 Competing Explanations.* J Quant Criminol, 1994. **10**(3): p. 189-212.
141. Lipsitt, PD, Buka, SL, and Lipsitt, LP, *Early Intelligence Scores and Subsequent Delinquency - a Prospective-Study.* Am J Fam Ther, 1990. **18**(2): p. 197-208.
142. Stattin, H and Klackenberg-Larsson, I, *Early language and intelligence development and their relationship to future criminal behavior.* J Abnorm Psychol, 1993. **102**(3): p. 369-78.
143. Fergusson, DM, Horwood, LJ, and Ridder, EM, *Show me the child at seven: the consequences of conduct problems in childhood for psychosocial functioning in adulthood.* J Child Psychol Psychiatry, 2005. **46**(8): p. 837-49.
144. Cole, SR and Hernan, MA, *Fallibility in estimating direct effects.* Int J Epidemiol, 2002. **31**(1): p. 163-65.
145. Moffitt, TE, *The Neuropsychology of Juvenile-Delinquency - a Critical-Review.* Crime Justice, 1990. **12**: p. 99-169.
146. Jepsen, JRM, Fagerlund, B, and Mortensen, EL, *Do Attention Deficits Influence IQ Assessment in Children and Adolescents With ADHD?* J Atten Disord, 2009. **12**(6): p. 551-62.
147. Pearl, J, *Causality: models, reasoning, and inference,* 2000, Cambridge: Cambridge University Press.
148. Greenland, S, Pearl, J, and Robins, JM, *Causal diagrams for epidemiologic research.* Epidemiology, 1999. **10**(1): p. 37-48.
149. Cole, SR, et al., *Illustrating bias due to conditioning on a collider.* Int J Epidemiol, 2010. **39**(2): p. 417-20.
150. Janszky, I, Ahlbom, A, and Svensson, AC, *The Janus face of statistical adjustment: confounders versus colliders.* Eur J Epidemiol, 2010. **25**(6): p. 361-3.
151. Moffitt, TE, *Genetic and environmental influences on antisocial behaviors: evidence from behavioral-genetic research.* Adv Genet, 2005. **55**: p. 41-104.
152. Carlin, JB, et al., *Regression models for twin studies: a critical review.* Int J Epidemiol, 2005. **34**(5): p. 1089-99.
153. Begg, MD and Parides, MK, *Separation of individual-level and cluster-level covariate effects in regression analysis of correlated data.* Stat Med, 2003. **22**(16): p. 2591-602.
154. Dwyer, T and Blizzard, L, *A discussion of some statistical methods for separating within-pair associations from associations among all twins in research on fetal origins of disease.* Paediatr Perinat Epidemiol, 2005. **19 Suppl 1**: p. 48-53.
155. Donovan, SJ and Susser, E, *Commentary: Advent of sibling designs.* Int J Epidemiol, 2011. **40**(2): p. 345-9.
156. Dwyer, T, Morley, R, and Blizzard, L, *Twins and fetal origins hypothesis: within-pair analyses.* Lancet, 2002. **359**(9324): p. 2205-6.
157. Groen-Blokhuis, MM, et al., *Evidence for a causal association of low birth weight and attention problems.* J Am Acad Child Adolesc Psychiatry, 2011. **50**(12): p. 1247-54 e2.
158. Ijzerman, RG, Boomsma, DI, and Stehouwer, CD, *Twins and the fetal origins hypothesis.* Lancet, 2002. **360**(9350): p. 2075.
159. Sjölander, A, Frisell, T, and Öberg, S, *Causal interpretation of between-within models for twin research.* Epidemiologic Methods, 2012: In Press.
160. Griliches, Z, *Sibling Models and Data in Economics - Beginnings of a Survey.* J Polit Econ, 1979. **87**(5): p. S37-S64.

161. Bound, J and Solon, G, *Double trouble: on the value of twins-based estimation of the return to schooling.* Econ Educ Rev, 1999. **18**(2): p. 169-82.
162. *Finishing the euchromatic sequence of the human genome.* Nature, 2004. **431**(7011): p. 931-45.
163. Lewin, B, *Genes 9*2008: Jones and Bartlett Publishers.
164. Scally, A, et al., *Insights into hominid evolution from the gorilla genome sequence.* Nature, 2012. **483**(7388): p. 169-75.
165. Chimpanzee Sequencing and Analysis Consortium. *Initial sequence of the chimpanzee genome and comparison with the human genome.* Nature, 2005. **437**(7055): p. 69-87.
166. Kehrer-Sawatzki, H and Cooper, DN, *Understanding the recent evolution of the human genome: insights from human-chimpanzee genome comparisons.* Hum Mutat, 2007. **28**(2): p. 99-130.
167. Fischer, R, *The correlation between relatives on the supposition of Mendelian inheritance.* Transactions of the Royal Society of Edinburgh, 1918. **52**: p. 399-433.
168. Falconer, DS and Mackay, TFC, *Introdution to Quantitative Genetics*. 4th ed1996: Pearson Education Limited.
169. Visscher, PM, Hill, WG, and Wray, NR, *Heritability in the genomics era - concepts and misconceptions.* Nat Rev Genet, 2008. **9**(4): p. 255-66.
170. Todorov, AA and Suarez, BK, *Genetic Liability Model*, in *Encyclopedia of Biostatistics*2005, John Wiley & Sons, Ltd.
171. Divgi, DR, *Calculation of the tetrachoric correlation coefficient.* Psychometrika, 1979. **44**(2): p. 169-72.
172. Pawitan, Y, et al., *Estimation of genetic and environmental factors for binary traits using family data.* Stat Med, 2004. **23**(3): p. 449-65.
173. Hopper, JL, *The Epidemiology of Genetic Epidemiology.* Acta Genet Med Gemel, 1992. **41**(4): p. 261-73.
174. Visscher, PM, et al., *Assumption-Free Estimation of Heritability from Genome-Wide Identity-by-Descent Sharing between Full Siblings.* PLoS Genet, 2006. **2**(3): p. e41.
175. Visscher, PM, et al., *Genome partitioning of genetic variation for height from 11,214 sibling pairs.* Am J Hum Genet, 2007. **81**(5): p. 1104-10.
176. Yang, J, et al., *GCTA: a tool for genome-wide complex trait analysis.* Am J Hum Genet, 2011. **88**(1): p. 76-82.
177. Davies, G, et al., *Genome-wide association studies establish that human intelligence is highly heritable and polygenic.* Mol Psychiatr, 2011. **16**(10): p. 996-1005.
178. Yang, J, et al., *Genome partitioning of genetic variation for complex traits using common SNPs.* Nat Genet, 2011. **43**(6): p. 519-25.
179. Kendler, KS, *"A gene for...": the nature of gene action in psychiatric disorders.* Am J Psychiatry, 2005. **162**(7): p. 1243-52.
180. Gunter, TD, Vaughn, MG, and Philibert, RA, *Behavioral Genetics in Antisocial Spectrum Disorders and Psychopathy: A Review of the Recent Literature.* Behav Sci Law, 2010. **28**(2): p. 148-73.
181. Lango Allen, H, et al., *Hundreds of variants clustered in genomic loci and biological pathways affect human height.* Nature, 2010. **467**(7317): p. 832-38.
182. Craig, IW and Halton, KE, *Genetics of human aggressive behaviour.* Hum Genet, 2009. **126**(1): p. 101-13.
183. *Genealogy by Genetics*. URL: http://www.thewarriorgene.com/. Accessed on May 1st 2012.
184. *MAOA Gene in GeneCards*. Available from: http://www.genecards.org/cgi-bin/carddisp.pl?gene=MAOA. Accessed on April 26th 2012
185. Caspi, A, et al., *Role of genotype in the cycle of violence in maltreated children.* Science, 2002. **297**(5582): p. 851-4.
186. Kim-Cohen, J, et al., *MAOA, maltreatment, and gene-environment interaction predicting children's mental health: new evidence and a meta-analysis.* Mol Psychiatr, 2006. **11**(10): p. 903-13.
187. Caspi, A, et al., *Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene.* Science, 2003. **301**(5631): p. 386-9.
188. Risch, N, et al., *Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis.* Jama, 2009. **301**(23): p. 2462-71.
189. Ficks, C and Waldman, I, *Candidate genes for aggression and antisocial behavior: A meta-analysis of the 5-HTTLPR and MAOA-uVNTR.* Behav Genet, 2011. **41**(6): p. 906-06.
190. The National Board of Health and Welfare, *Adoption – Handbok för socialtjänsten* 2008.

191. Lichtenstein, P, et al., *The Swedish Twin Registry in the third millennium: an update.* Twin Res Hum Genet, 2006. **9**(6): p. 875-82.
192. Lichtenstein, P, et al., *The Swedish Twin Registry: a unique resource for clinical, epidemiological and genetic studies.* J Intern Med, 2002. **252**(3): p. 184-205.
193. Swedish National Court Administration, *Kvalitet och förtroende. Rapport 1987:8*, 1987, Domstolsverket: Jönköping.
194. Swedish National Court Administration, *Överklagande och ändring. Rapport 1995:3.*, 1995, Domstolsverket: Jönköping.
195. Ross, A, *De värnpliktigas prestationsförmåga vid inskrivningsprövningar i Sverige 1969-1979: bakgrundsvariabler relaterade till intellektuell prestationsförmåga : en beskrivning av data från intervjuunderlag och exempel på modeller*. 1988: Pedagogiska institutionen, Umeå univ.
196. Carlstedt, B and Mardberg, B, *Construct-Validity of the Swedish Enlistment Battery.* Scand J Psychol, 1993. **34**(4): p. 353-62.
197. Mardberg, B and Carlstedt, B, *Swedish Enlistment Battery (SEB): Construct validity and latent variable estimation of cognitive abilities by the CAT-SEB.* Int J Select Assess, 1998. **6**(2): p. 107-14.
198. Langholz, B, *Case–Control Study, Nested*, in *Encyclopedia of Biostatistics.* 2005, John Wiley & Sons, Ltd.
199. D, C and M, H, *Statistical Models in Epidemiology.* 2002, Oxford: Oxford University Press.
200. Rothman, KJ, Greenland, S, and Lash, TL, *Modern Epidemiology*. Third ed2008: Lippincott Williams & Wilkins.
201. Pickles, A, *Generalized Estimating Equations*, in *Encyclopedia of Biostatistics.* 2005, John Wiley & Sons, Ltd.
202. Fitzmaurice, GM, Laird, NM, and Ware, JH, *Applied Longitudinal Analysis.* 2011: John Wiley & Sons.
203. Genz, A, *Numerical computation of multivariate normal probabilities.* Journal of Computational and Graphical Statistics, 1992. **1**: p. 141-49.
204. Pawitan, Y, *In All Likelihood: Statistical Modelling and Inference Using Likelihood.* 2001: Clarendon Press.
205. Sjölander, A, et al., *Analysis of 1:1 matched cohort studies and twin studies, with binary exposures and binary outcomes.* Stat Med, 2012(In press).
206. Öberg, S, et al., *Birth Weight-Breast Cancer Revisited: Is the Association Confounded by Familial Factors?* Cancer Epidem Biomar, 2009. **18**(9): p. 2447-52.
207. Greenland, S, *Collapsibility*, in *Encyclopedia of Biostatistics.* 2005, John Wiley & Sons, Ltd.
208. Lin, DY and Wei, LJ, *The robust inference for the Cox proportional hazards model.* J Am Stat Assoc, 1989. **84**(408): p. 1074-78.
209. Leu, M, Czene, K, and Reilly, M, *The impact of truncation and missing family links in population-based registers on familial risk estimates.* Am J Epidemiol, 2007. **166**(12): p. 1461-7.
210. Bellis, MA, et al., *Measuring paternal discrepancy and its public health consequences.* J Epidemiol Community Health, 2005. **59**(9): p. 749-54.
211. Carey, G, *Twin immitation for antisocial behavior: Implications for genetic and family environment research.* J Abnorm Psychol, 1992. **101**(1): p. 18-25.
212. Patrick, CJ and Vaidyanathan, U, *Coming to grips with the cycle of violence.* Psychol Med, 2011. **41**(1): p. 41-5.
213. McGue, M, et al., *The environments of adopted and non-adopted youth: evidence on range restriction from the Sibling Interaction and Behavior Study (SIBS).* Behav Genet, 2007. **37**(3): p. 449-62.
214. Aulchenko, YS, et al., *Predicting human height by Victorian and genomic methods.* Eur J Hum Genet, 2009. **17**(8): p. 1070-5.
215. Stein, JL, et al., *Identification of common variants associated with human hippocampal and intracranial volumes.* Nat Genet, 2012. **advance online publication**.
216. *HMGA2 Gene in GeneCards*. Available from: http://www.genecards.org/cgi-bin/carddisp.pl?gene=HMGA2. Accessed on April 26th 2012.
217. Bergen, SE and Petryshen, TL, *Genome-wide association studies of schizophrenia: does bigger lead to better results?* Curr Opin Psychiatr, 2012. **25**(2): p. 76-82.
218. Lee, KW, et al., *Genome wide association studies (GWAS) and copy number variation (CNV) studies of the major psychoses: what have we learnt?* Neurosci Biobehav Rev, 2012. **36**(1): p. 556-71.