Institute for Molecular Medicine Finland (FIMM),
HiLIFE,
and
Faculty of Medicine
Doctoral Program in Population Health
University of Helsinki

Associations with childhood aggressive behavior using multiple raters: co-occurrence, antisocial personality disorder prediction, and biomarkers

Alyce M Whipp

ACADEMIC DISSERTATION

Doctoral thesis, to be presented for public examination with the permission of the Faculty of Medicine of the University of Helsinki, in Lecture Hall 2, Haartman Institute, Haartmaninkatu 3, on 15 October 2021 at noon.

Helsinki, Finland 2021

SUPERVISORS:

Docent Tellervo Korhonen, PhD
Institute for Molecular Medicine Finland (FIMM)
HiLIFE
University of Helsinki, Finland

Docent Eero Vuoksimaa, PhD
Institute for Molecular Medicine Finland (FIMM)
HiLIFE
University of Helsinki, Finland

REVIEWERS:

Professor Pirkko Riipinen, MD PhD
Department of Psychiatry
University of Oulu, Finland

Professor Tommi Tolmunen, MD PhD
Department of Adolescent Psychiatry
University of Eastern Finland, Finland

OPPONENT:

Professor Essi Viding, PhD
Division of Psychology and Language Sciences
University College London, UK

The Faculty of Medicine uses the Urkund system (plagiarism recognition) to examine all doctoral dissertations.

Dissertationes Scholae Doctoralis Ad Sanitatem Investigandam Universitatis Helsinkiensis No.47/2021

Cover layout by Anita Tienhaara

ISBN 978-951-51-7493-2 (paperback)
ISBN 978-951-51-7494-9 (PDF)
ISSN 2342-3161 (print)
ISSN 2342-317X (online)

http://ethesis.helsinki.fi/

Unigrafia, Helsinki 2021

To my dear matriarch figures:
Susan, Mary Ann, Ursula, and Rena

"Believe those who seek the truth, doubt those who found it" -André Gide

ABSTRACT

Problems of aggressive behavior affect as many as one in every six children and are associated with negative outcomes for not only the individual themselves, but also their family, friends, and community. Aggressive behavior includes a variety of different behaviors (e.g., yelling, hitting, bullying) and has been notoriously difficult to treat. In assessing aggression, researchers and clinicians have repeatedly been recommended to utilize reports from multiple informants (e.g., parents, teachers, the child him/herself) in order to obtain the most complete picture of the problem. In studying aggressive behavior, early research was initially focused on males only and severe outcomes, but now investigates gender similarities and differences and a broad range of behaviors related to aggression. While much has been learned from research thus far, new biological mechanisms and detailed phenotypic information are still important to continue clarifying the heterogenous nature of aggression and to improve ideas for personalized treatment. Thus, this thesis aimed to contribute to those efforts.

Study I and II comprised cohorts of thousands of children from Finland, the Netherlands, and the UK (Study I also included Swedish cohorts) with ratings of aggressive behavior, other externalizing behaviors (e.g., hyperactivity), internalizing problems (e.g., depression) and prosocial behavior from parents, teachers, and the children themselves. Results showed that aggressive behavior (as rated by all raters) often co-occurred with other externalizing behaviors and low prosocial behavior, and also co-occurred with internalizing problems but not as often. Patterns were similar across cohorts and genders, however, parents indicated more co-occurrence with internalizing problems and less co-occurrence with other externalizing behaviors than teachers did.

Study III utilized a subset of the FinnTwin12 cohort (N=1347) with longitudinal data on aggressive behavior in adolescence from parents, teachers, children themselves, and from their co-twin to investigate prediction of antisocial personality disorder (ASPD) in young adulthood. Notably, teacher and self ratings were able to predict ASPD, both in separate models and when both ratings were in the same model. Additionally, the direct aggression subtype (e.g., hitting, yelling) was able to predict ASPD well, for both genders. Furthermore, when the co-occurring influence of hyperactivity was removed from the aggression ratings (using a residual aggression variable with hyperactivity co-occurrence removed), aggressive behavior was still able to predict ASPD.

Study IV utilized a subset of the FinnTwin12 cohort (N=725) with plasma samples in young adulthood and aggressive behavior measures in

adolescence from parents, teachers, children themselves, and from their co-twin to search for new biomarkers of aggression. Using a panel of 11 metabolites processed using nuclear magnetic resonance, the ketone body 3-hydroxybutyrate was found to be negatively associated with aggressive behavior in initial analyses. In more detailed modeling, nearly all raters of aggression showed the same trend with 3-hydroxybutyrate, including in fully adjusted models. In a model including both teacher and self ratings, 3-hydroxybutyrate was significantly associated with both aggressive behavior ratings. A replication dataset of young adult Dutch twins (N=960) showed support for the association found in FinnTwin12, however, the issue of whether there are gender differences of the association of 3-hydroxybutyrate with aggressive behavior remains to be clarified by future research.

These findings help to clarify the co-occurrence of aggressive behavior with other behaviors across raters and countries, to show how common the co-occurrence is and that it should be taken into consideration when studying aggressive behavior, including from (epi)genetic or biological perspectives. Additionally, aggressive behavior, in particular direct aggression, can inform future ASPD risk, and obtaining behavior data from teachers and the child are of high importance. Furthermore, the new association of 3-hydroxybutyrate with aggressive behavior suggests new biological pathways to investigate to improve our understanding of aggressive behavior, including potential treatments. This thesis provides refinements to the aggressive behavior phenotype, new avenues for aggression biology investigations, and ideas for where to improve or personalize treatment options.

TIIVISTELMÄ

Aggressiiviseen käyttäytymiseen liittyvät ongelmat vaikuttavat jopa joka kuudenteen lapseen, ja niihin liittyy kielteisiä seurauksia niin henkilölle itselleen kuin myös hänen läheisilleen.. Aggressiivinen käyttäytyminen käsittää monenlaista ilmenemismuotoja (esim. huutaminen, lyöminen), ja sitä on tunnetusti vaikea hoitaa. Usein suositellaan, että aggressiivisuuden arvioinnissa kerättäisiin tietoa useista eri lähteistä (esim. vanhemmilta ja opettajilta), jotta ongelmasta saataisiin mahdollisimman kattava kuva. Alkuvaiheessa aggressiivisen käyttäytymisen tutkimuksessa keskityttiin lähinnä miehiin ja vaikeisiin aggressiivisen käytöksen ilmenemismuotoihin. Nykypäivänä tutkitaan myös sukupuolten välisiä yhtäläisyyksiä tai eroja, ja aggressiivisuuteen liittyvä käyttäytyminen ymmärretään laajempana jatkumona. Vaikka tähänastisesta tutkimuksesta on kertynyt paljon tietämystä, on edelleen tärkeää tutkia uusia biologisia mekanismeja ja erilaisia fenotyyppejä aggression esiintymismuotoja/ilmiasuja heterogeenisen luonteen selvittämiseksi ja uudenlaisten yksilöllisten hoitomuotojen kehittämiseksi. Tämän väitöskirjan tavoitteena oli tuottaa lisätietoa näistä aiheista.

Osatyöt I ja II osoittivat, että aggressiivista käyttäytymistä esiintyi usein yhdessä muiden eksternalisoivien käyttäytymisen ilmenemismuotojen ja vähäisen prososiaalisen käyttäytymisen kanssa. Sitä esiintyi myös yhdessä internalisoivien ongelmien kanssa mutta harvemmin. Nämä havainnot toistuivat samankaltaisina niin pojilla kuin tytöilläkin sekä eri maissakin, mutta vanhempien raportoimana aggressiivista käyttäytymistä ilmoitettiin esiintyvän enemmän yhdessä internalisoivien ongelmien kanssa ja vähemmän yhdessä muiden eksternalisoivien käyttäytymismallien kanssa kuin opettajien raportoimana.

Osatyö III osoitti, että lapsuuden aggressiivinen käyttäytyminen (erityisesti opettajien ja lapsen itsensä arvioimana) voi ennustaa epäsosiaalista persoonallisuushäiriötä nuorena aikuisuisena. Lisäksi suoran aggression alatyyppi (esim. lyöminen, huutaminen) ennusti epäsosiaalista persoonallisuushäiriötä hyvin molemmilla sukupuolilla. Aggressiivinen käyttäytyminen ennusti edelleen epäsosiaalista persoonallisuushäiriötä vaikka lasten yliaktiivisuus otettiin huomioon.

Osatyö IV osoitti, että ketoaine 3-hydroksivoihappo oli negatiivisesti yhteydessä aggressiiviseen käyttäytymiseen. Yhteys havaittiin niin

Suomalaisessa kuin Hollantilaisessakin aineistossa. Jatkotutkimukselle jää kuitenkin selvitettäväksi kysymys siitä, voiko yhteyteen liittyä sukupuolieroja.

Tässä väitöskirjassa kuvaillaan yksityiskohtaisemmin aggressiivisen käyttäytymisen fenotyyppiä, esitellään uusia lähestymistapoja aggression biologisiin tutkimuksiin sekä tarjotaan uusia näkökulmia hoitovaihtoehtojen kehittämiseen ja yksilöllistämiseen. Tulokset selvensivät aggressiivisen käyttäytymisen ja muiden käyttäytymismuotojen välistä yhteyttä eri arvioijien raportoimana sekä eri maissa. Tulokset osoittavat, kuinka yleistä niiden samanaikainen esiintyminen on, ja että se olisi otettava huomioon. Aggressiivinen käyttäytyminen, erityisesti suora aggressio, voi myös kertoa kehittymässä olevan epäsosiaalisen persoonallisuushäiriön riskistä, ja on hyvin tärkeää hankkia käyttäytymistä koskevaa tietoa sekä opettajilta että lapselta itseltään. Lisäksi uusi löydös 3-hydroksivoihapon yhteydestä aggressiiviseen käyttäytymiseen viittaa uusiin biologisiin mekanismeihin. Niitä on tutkittava tarkemmin, jotta voimme ymmärtää aggressiivisesta käyttäytymistä paremmin ja kehittää häiritsevän aggressiivisen käyttäytymisen hoitoja.

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following articles which are referred to in the text by their Roman numeral**:**

I. Bartels M, Hendriks A, Mauri M, Krapohl E, **Whipp A**, Bolhuis K, Conde LC, Luningham L, Fung Ip H, Hagenbeek F, Roetman P, Gatej R, Lamers A, Nivard M, van Dongen J, Lu Y, Middeldorp C, van Beijsterveldt T, Vermeiren R, Hankemeijer T, Kluft C, Medland S, Lundstrom S, Rose R, Pulkkinen L, Vuoksimaa E, Korhonen T, Martin NG, Lubke G, Finkenauer C, Fanos V, Tiemeier H, Lichtenstein P, Plomin R, Kaprio J and Boomsma DI. 2018. Childhood aggression and the co-occurrence of behavioural and emotional problems: results across ages 3–16 years from multiple raters in six cohorts in the EU-ACTION project. European Child & Adolescent Psychiatry. 27(9): 1105-1121. doi: 10.1007/s00787-018-1169-1

II. **Whipp AM**, Vuoksimaa E, Bolhuis K, de Zeeuw EL, Korhonen T, Mauri M, Pulkkinen L, Rimfeld K, Rose RJ, van Beijsterveldt T CEM, Bartels M, Plomin R, Tiemeier H, Kaprio J, Boomsma DI. 2021. Teacher-rated aggression and co-occurring behaviors among schoolchildren: A comparison of four population-based European cohorts. *PLOS One*. https://doi.org/10.1371/journal.pone.0238667

III. **Whipp AM,** Korhonen T, Raevuori A, Heikkilä K, Pulkkinen L, Rose RJ, Kaprio J, Vuoksimaa E. 2019. Early adolescent aggression predicts antisocial personality disorder in young adults: a population-based study. *European Child & Adolescent Psychiatry*. 28(3), 341-350. doi:10.1007/s00787-018-1198-9

IV. **Whipp AM**, Vuoksimaa E, Korhonen T, Ligthart RSL, Pool R, Hagenbeek F, Bogl LH, Pulkkinen L, Rose RJ, Boomsma D, Kaprio J. 2021. Ketone 3-hydroxybutyrate as a biomarker of aggression. Scientific Reports. 11, 5813. doi: 10.1038/s41598-021-84635-6

ABBREVIATIONS

ADHD	Attention-deficit/hyperactivity disorder
ASEBA	Achenbach System of Empirical-Based Assessment
ASPD	Antisocial personality disorder
ASR	Adult Self-Report
A-TAC	Autism–Tics, ADHD and other Comorbidities inventory
AUC	Area under the curve
BMI	Body mass index
CATSS	Child and Adolescent Twin Study in Sweden
CBCL	Child Behavior Checklist
CD	Conduct disorder
CI	Confidence interval
DCB	Devereux Child Behavior
DNA	Deoxyribonucleic acid
DSM	Diagnostic and Statistical Manual of Mental Disorders
FemNAT-CD	Neurobiology and Treatment of Adolescent Female Conduct Disorder: The Central Role of Emotion Processing
FT12	FinnTwin12
GABA	Gamma-aminobutyric acid
GENR	Generation R
IED	Intermittent explosive disorder
MAO	Mono-amine oxidase
METS	Metabolic equivalent of task
MPNI	Multi-dimensional Peer Nomination Inventory
MS	Mass spectrometry
NMR	Nuclear magnetic resonance spectroscopy
NTR	Netherlands Twin Register
ODD	Oppositional defiant disorder
OR	Odds ratio
PATHS	Promoting Alternative Thinking Strategies
RNA	Ribonucleic acid
ROC	Receiver operating characteristic
SAAF	Strong African American Families
SDQ	Strengths and Difficulties Questionnaire
SSAGA	Semi-Structured Assessment for the Genetics of Alcoholism
TCHAD	Twin study of Child and Adolescent Development
TEDS	Twins Early Development Study
TRF	Teacher Report Form
YSR	Youth Self-Report

INTRODUCTION

Aggressive behavior is one of our most primal behaviors, to some extent we need it to survive. However, high levels of aggressive behavior are maladaptive in modern human society. Developmentally, most humans exhibit aggressive behavior as young children, developing the ability to better control their behavior around school age, when aggressive behavior has largely declined for the majority of individuals [1]. Children who enter school still showing high levels of aggression are at risk for developing aggressive behavior disorders. The lifetime prevalence for aggressive behavior disorders (e.g., conduct disorder, oppositional defiant disorder) ranges 7.9–13.9% for adolescent boys and 5.8–11.3% for adolescent girls, with a median age of onset of 11 years old [2]. Children as young as 7–9 years old with high levels of aggressive behavior have been shown to be at much greater risk of criminal behavior, substance (ab)use, mental health problems, risky sexual behavior, and intimate partner violence 25 years later [3]. However, aggressive behavior has been notoriously difficult to treat; we need a better understanding of the behavior in order to improve treatments.

Our understanding of aggressive behavior is complicated by the fact that it is heterogenous in nature. There are many subtypes of aggression (e.g., direct, indirect, proactive, reactive) and the behavior can be exhibited across all environments an individual interacts in, or only some. The more generalized the aggressive behavior (i.e., the behavior occurs in a broad range of environments instead of a specific one), the more serious the aggressive behavior problem, generally [4]. In order to capture the most complete scope of the aggressive behavior problem as possible, researchers and clinicians collect information from a number of informants: parents, the child him/herself (if they are old enough), teachers. Unfortunately, these multiple informants are not known to show high inter-rater agreement [5]. Taking these multiple informants into consideration for their overlapping and unique contributions to the breadth and depth of the aggressive behavior problems is important for our understanding.

In early research, aggressive behavior studies were focused on males only, and often on criminal or severe outcomes. However, over the years, female aggressive behavior has become better characterized and the study of aggressive behavior has broadened. Findings of gender differences in aggressive behavior indicate that while they do exist, the nuances are important and the differences are perhaps not as stark as was once imagined [6, 7]. Broadening our investigations of aggressive behavior has also shown that it is important to not only focus research efforts on specific manifestations of aggression (e.g., bullying or intimate partner violence), but it is also useful

to look at broader categories of behaviors of which aggression is a part (e.g., disruptive behaviors, externalizing behaviors). The examination of similarities and differences regarding gender, aggression subtypes, and fine-scale and broad-scale aggressive behavior categories has helped to explain more of the heterogeneousness of aggressive behavior, but there is still more yet to learn.

In recent decades, research has included (epi)genetics, neuroscience, and biomarkers in the search for new and better insights into aggressive behavior. It is hoped that these areas of research can improve the ability to create effective treatments, however, for them to provide the most useful results they need to utilize well-defined phenotypes of aggression. Thus, as more advanced approaches are being applied in aggression research, we still need to make sure the characterization of aggressive behavior that is used is as complete and detailed as possible. To these ends, the main objective of this thesis is to improve our understanding and characterization of the aggression phenotype. Two studies look across multiple European cohorts using multiple informants to characterize what behaviors and emotional problems co-occur with aggressive behavior. Another study examines whether adolescent aggressive behavior (alone or including the influence of co-occuring behaviors) can predict a young adult personality disorder; this investigation also looks at subtypes of aggression regarding the prediction. The final study uses a hypothesis-free approach to search for associations of aggressive behavior with a small panel of metabolites in order to generate new ideas for future research and improved treatment.

REVIEW OF THE LITERATURE

As the king said in Lewis Carroll's *Alice's Adventures in Wonderland*, "Begin at the beginning", but regarding the study of human aggression, it can be difficult to determine where exactly that beginning should be. Do we begin in the individual body and mind, describing where aggressive behavior comes from biologically? Or do we go broader and begin by discussing theories of why humans aggress or where aggression 'starts' and develops over a human's lifetime? Or do we start still further afield and begin at aggression's primeval roots in the animal kingdom, why and where the behavior came about and why it persisted evolutionarily? A simple though incomplete response is that it depends on the scope of the discussion. Since my research involves only human aggressive behavior and uses epidemiological approaches, this will serve to limit the scope to only some of the abovementioned areas of aggression research. Regardless, because the aggression phenotype has piqued scientific interest for many decades and from many different fields and perspectives, it would be wise to begin with defining the term, and its associated concepts.

2.1 Introduction to Aggressive Behavior

This initial section introduces the definitions of aggressive behavior, its subtypes, and psychosocial theories of its origin. I also briefly highlight some of the biology of aggression, before describing typical human development of aggressive behavior and its stability over a human lifetime, and discuss the predictors and outcomes of problem levels of aggressive behavior. Throughout, I clarify any differences or similarities in aggressive behavior patterns by sex/gender. In this thesis, 'sex' refers to biological distinctions between males and females, while 'gender' refers to the social constructs and roles that typically distinguish men/boys from women/girls. Sometimes it is unclear if we are measuring one or the other or both, but an attempt to adhere to these distinctions is made.

2.1.1. Definition(s)

As Richard Tremblay, a well-known developmental psychologist focusing on childhood aggression, pointed out in his review of 20th century research on the development of aggressive behavior during childhood [1], one could argue that a major problem with studying this topic is the lack of agreement on a definition of aggression. There seem to be at least two main aspects involved in the murky definitions: the intent to harm, and the social (un)desirableness

of the action. According to Berkowitz [8] and others, aggression is any behavior with the *intent* to inflict physical or mental harm or injury on another, whereas Loeber and Stouthamer-Loeber [9] remove the word 'intent' and define it simply as 'those acts that inflict bodily or mental harm on others' (p. 242). Alternatively, Bandura [10] and others rather define aggression as any harmful behavior which is in violation of social norms. One complication of definitions incorporating social norms is that observers of the potential aggression have to make a judgment call, and this often involves their opinion of whether the aggressor intended to hurt the victim. Thus, perhaps we can argue that much of the definition debate revolves around this 'intent to harm' concept, which, as we will see in later discussion of the subtypes of aggression (section 2.1.2), gets perhaps murkier still regarding, for example, reactive aggression. Therefore, perhaps the more broad definition offered by Loeber and Stouthamer-Loeber [9], which is silent on intent, is the more inclusive choice in this broad literature review context.

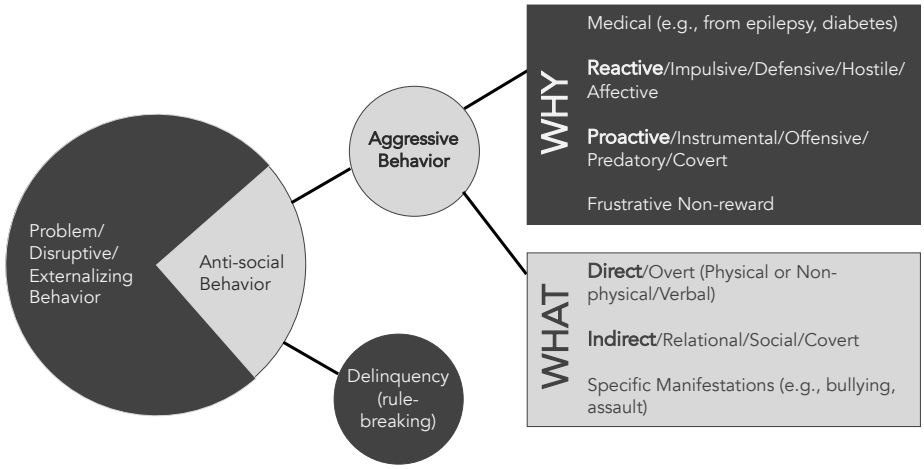
Further semantics would assist here to clarify other aspects of aggression terminology. In a meta-analytic review of personality and aggressive behavior [11], the term *aggression* is identified broadly and with the intent aspect mentioned above. However, they then go on to distinguish *aggressive behavior* (acts of aggression) from *trait aggressiveness*. Trait aggressiveness refers, in personality psychology literature, to 'people who are prone to hostile cognitions and angry affect as well as a readiness to engage in physical and verbal aggression' (p. 752) [11] [see also 12]. Some of the literature also distinguishes between trait and state aggression, with trait variables capturing more typical or stable aggressive aspects, whereas state variables capture more temporary aggressive aspects. Although one's personality (our typical patterns of behavior, thought, and feelings) is an important aspect in how or why we behave the way we do, personality and trait aggressiveness are their own distinct field of research, and one that is not the focus of the research done here, thus I will maintain focus on *aggressive behavior* only.

The remaining aspect to define regarding aggressive behavior research is how or if we will limit the broadness of the manifestations of the aggressive behaviors we study. Some researchers zoom in to focus on specific manifestations such as bullying, homicide, or domestic violence. Those bodies of research of course help us to better understand those specific (often severe) behaviors, but it does not necessarily help us make broader connections and understand overarching similarities across all aggressive behaviors. In contrast, we can incorporate aggressive behavior with many other seemingly similar behavioral problems into an ever-broadening array of categories, possibly all the way up to simply 'socially unacceptable'. Tremblay [1], however, acknowledges that these perhaps too-broad categorizations of aggressive behavior complicate our study of it. Researchers have used, for example, such broader categories as 'antisocial behavior', 'externalizing

behaviors', 'disruptive behaviors', 'problem behaviors' and 'conduct problems' to incorporate not only aggressive behavior but also other related behaviors and problems. For the purposes of this thesis, we aim to focus only on aggressive behaviors when we use said term, and to limit our inclusion of the broader categories whenever possible (although for some studies we specifically include other externalizing behaviors, their distinctions from aggressive behavior are made clear; further discussion on other externalizing behaviors occurs in section 2.2.1).

To summarize, the definition distinctions described above and further discussion about aggression subtypes (section 2.1.2) can become quite confusing in this literature. Therefore, an attempt by me to try to conceptually distinguish and relate some of these elements of the aggressive behavior literature was made in the form of Figure 1. This diagram clarifies my understanding of how all these elements relate to each other, based on my review of the literature. For example, the broader aggression-related categories mentioned above are indicated in the circle on the left in Figure 1. Furthermore, we see the term 'antisocial behavior' has been broken into 'aggressive behavior' and 'delinquency' categories ('delinquency' has also been called 'rule-breaking behavior' or 'non-aggressive antisocial behavior' in the literature). This separation appears well supported by research on the etiology of antisocial behavior and externalizing problems [see for example 13, 14]. This thesis mainly deals with the bolded 'Aggressive Behavior' and 'Direct' and 'Indirect' concepts of Figure 1, but all other aspects will be described in the following section 2.1.2.

Figure 1. Diagram of the terminology and concepts surrounding aggressive behavior



References used to compile this diagram: [1, 7, 13–21, 24]

2.1.2. Subtypes of Aggressive Behavior

Now that the general definitions of aggressive behavior and related terms have been clarified, it is important to define and understand the different ways researchers have subcategorized and subtyped aggressive behavior, and how these subtypes help us better understand aggressive behavior.

The initial important distinction regarding aggressive behavior subtypes is to introduce the fairly well agreed upon dimensions of aggressive behavior, what Little et al. [15] call the 'why' and the 'what' (see Figure 1). The 'why' refers to the motivation behind the aggressive action: 'Why did they do it? What caused them to behave that way?', while the 'what' refers to the type of aggressive action the person performed: 'What did they do? In what way did they aggress against the victim?' Each of these 'why' and 'what' categories can be further divided into contrasting aspects. First, however, Turgay [16] reminds us that any medical 'why' that can cause the aggressive behavior must be ruled out, since we know epilepsy, dementia, brain tumors, diabetes, and other conditions can cause aggressive behavior in patients.

The 'why' dimension can be decomposed into proactive and reactive aggression subtypes (see Figure 1)[17-19]. Proactive aggression is a goal-oriented, planned form of the behavior, and is sometimes also called 'instrumental', 'offensive', 'predatory' (in animal studies), or 'covert'. In contrast, reactive aggression is an impulsive, reactionary form of the behavior often arising from anger or frustration, and is sometimes called 'impulsive', 'defensive', 'hostile', or 'affective'. Additionally, based on the National Institute for Mental Health's Research Domain Criteria Initiative, this 'why' dimension might also be considered to have a 'frustrative non-reward' subtype [20]. This occurs after repeated, failed attempts to obtain something desired; however, since it also seems to share some elements of reactive aggression and is not well-studied, I will hereafter focus on the two main elements in the 'why' dimension: proactive and reactive.

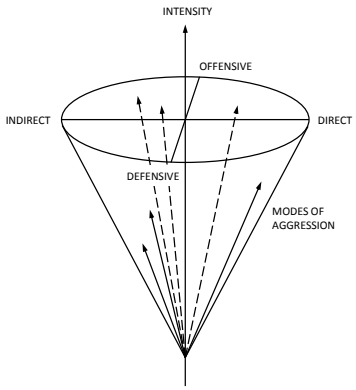
Proactive and reactive aggression are highly correlated (r ranging: 0.68–0.87 based on questionnaire data), however, they are considered to be distinct entities [17-19]. They are said to have different origins; proactive is proposed to arise from a history of positive reinforcement for aggression in one's social environment [10], while reactive is proposed to arise from one's temperament/personality and/or possibly traumas in the childhood environment [21]. Furthermore, reactive aggression appears to be more strongly associated with attention and impulsivity problems [22], compared to proactive aggression which is more strongly associated with psychopathy [22]. Additionally, there do not appear to be strong gender differences regarding these two subtypes [6, 23]. Lastly, because these two subtypes rely to some extent on knowing the 'intent to harm' (or impulse to harm in the case

of reactive aggression) of the aggressor, they are more challenging to study objectively.

Getting back to the 'what' dimension of aggression, this dimension can be decomposed into two main categories: direct and indirect (see Figure 1). Although strongly intercorrelated (r=0.76), they are still regarded at distinct types due to differential associations with maladjustment [7]. Direct aggression is visible or overt actions; these can be further categorized as physical (e.g., hitting, kicking, pushing) or non-physical/verbal (e.g., yelling, name-calling). In contrast, indirect aggression is a more covert action and has also been called 'relational' or 'social' aggression; in general, this subtype entails spreading rumors or socially ostrasizing others. There has been debate regarding whether the alternative terminologies for indirect aggression are synonyms or separate entities, but they are generally considered to be similar enough to group together [24]. Overall, there is substantially less research done on indirect, compared to direct, aggression, perhaps due to early findings suggesting males mostly physically/directly aggressed, and females mostly indirectly aggressed [25-28]. However, it has now generally been shown that males and females who are aggressive utilize both direct and indirect modes of aggression [7, 29].

In reality, each manifestation of aggression, and there are thousands of them, involves varying degrees of these different dimensions of aggressive behavior. In 1969, Finnish psychology researcher Lea Pulkkinen presented a pioneering visualization of a dimensional model of aggressive behavior combining the abovementioned dimensions (Figure 2)[19, 30]. She placed the 'why' dimensions (i.e., defensive/reactive and offensive/proactive) opposite each other, and the 'what' dimensions (i.e., direct and indirect) opposite each other, creating a circular shape. The model additionally takes the intensity of the action into consideration, creating an overall cone-like shape to the model. This conceptualization reminds us that although we can focus on one or a few specific manifestations of aggressive behavior, it is also important to keep the broader picture of aggressive behavior in mind, to have a more complete understanding of the behavior. However, there is a balancing act, since we know that these different subtypes have somewhat different determinants, lumping aggression into too broad a category can reduce our ability to investigate it effectively [1].

Figure 2. Lea Pulkkinen's dimensional model of aggression



Source: Pitkänen (1969, Figure 1, p. 29). Reproduced with permission.

2.1.3. Theories of Aggressive Behavior

There are numerous theories of the nature and origin of aggression through different lenses [31], including evolutionary, neurological, and psychological development. Because these vary in scope and perspective, it is most useful to focus on those that provide the best framework for understanding the nature and level of aggressive behavior one studies. As this thesis mostly focuses on human, particularly child/adolescent, aggressive behavior from a psychopathological perspective, I will focus on theories that explain psychosocial aspects of aggressive behavior that can lead to abnormal or maladaptive problems. However, in section 2.1.4, I also describe some biological pathways/mechanisms involved in aggression.

The three main theories of aggressive behavior discussed here are: Bandura's social learning theory, Berkowitz's frustration–aggression theory, and Crick & Dodge's social information processing theory. Additionally, the two main aggression theories that attempt to explain sex differences in aggression are described.

Bandura's social learning theory [10] suggests that people learn aggressive responses in the same way they learn other complex social behaviors, by direct experience or observing others. At an individual level, this suggests that having good (or bad) role models in one's life will thus have an important effect on one's behavior. At a societal/cultural level, observed gender roles can 'teach' men and women what kind of behaviors are expected or preferred. This theory is particularly useful in understanding and explaining proactive aggression, especially the aspect of reward resulting from certain behaviors.

Berkowitz's frustration–aggression theory [8] suggests that negative affect arises from aversive events such as frustrations, perceived or actual threats, and/or unpleasant noises, temperatures, or smells. The resulting negative affect then stimulates certain thoughts, memories, reactions, and physiological responses related to both fight and flight responses. Fight-related responses produce feelings of anger; flight-related responses produce feelings of fear. Additionally, during aversive events, certain cues become linked in one's memory with not only the event, but with the resulting cognitive and emotional responses triggered by the event. Thus, responses become 'conditioned' or 'learned'. This theory is particularly useful in understanding and explaining reactive aggression.

The social information processing theory of Crick and Dodge [32] blends aspects of Bandura's and Berkowitz's theories and describes steps that children take during social information processing. Linked in a repeating cycle, the steps include (a) interpreting social cues, (b) creating possible solutions to perceived problems, (c) evaluating one's own ability and efficacy to manage and perform the solutions, (d) evaluating the likelihood that the proposed solutions will result in a positive outcome, and (e) clarifying one's goals in the social situation. Children characterized as exhibiting the reactive subtype of aggression have been shown to perceive hostile intent (in step (a) from above) in ambiguous situations, but can generally correctly perceive the cues of clearly benign and clearly hostile situations [32]. Additionally, this hostile attributional style is more likely to occur in children who experience peer rejection or harsh parenting [32]. Furthermore, reactive aggression has been more strongly associated with creating aggressive responses to situations of conflict (in step (b) from above) than proactive aggression [33]. In contrast, proactive aggression is associated with problems in later steps of information processing (steps (d) and (e))[17].

The main aggression theories that attempt to explain sex differences roughly correspond to the two sides of the classic nature–nurture debate: sexual selection theory and social role theory (which relates to Bandura's social learning theory), respectively. Briefly, sexual selection theory has an evolutionary explanation (i.e., is not exclusively a human-based theory). Males experience competition for access to reproduction (i.e., females) due to differences in parental investment; thus, males are more likely to be aggressive. Social role theory suggests that gender roles in a society allow or encourage differing amounts of aggressive or competitive behavior in men and women [34], suggesting males are more likely to be physically aggressive. Neither theory alone explains the sex/gender differences/similarities seen in human aggression, but Wölfer and Hewstone [35] suggest that sexual selection theory explains intrasex aggression well (i.e., within a sex, particularly males), while social role theory explains intersex aggression well (i.e., between the sexes, particularly male-on-female aggression), and this

seems to hold true across varying levels of ethnic diversity. However, in contrast to social role theory expectations, Nivette et al. [36] showed that across low- and middle-income countries, as gender inequality increased, gender differences in physical aggression decreased; they also found no evidence of sexual selection theory as an explanation for the gender differences.

2.1.4. Biology of Aggressive Behavior

While the psychosocial theories of aggressive behavior are important in understanding how people process emotions and information leading to behaving or not behaving aggressively, it is also important to go down to the biological level to see what biochemical pathways and mechanisms are contributing with our psychosocial processing. The biology of aggression is an enormous area of research, often studied in animals, that spans neurocognitive research, biomarkers/biochemistry, (epi)genetics, and more. Since this thesis does not utilize any laborotory or bioinformatic techniques, and instead focuses on biomarkers from an epidemiological standpoint, I will only briefly focus on the main biochemical pathways in aggression: neurotransmitters, hormones, and inflammatory biomarkers.

The three main neurotransmitters (and their pathways) involved in aggression are serotonin, dopamine, and GABA (gamma-aminobutyric acid). Dopamine contributes to initiating aggressive behavior, while serotonin contributes to inhibiting the behavior [37]. Reduced activity of the serotonin pathway, which regulates the dopamine pathway, leads to elevated activity of the dopamine pathway [38]. The molecule mono-amine oxidase (MAO), which breaks down serotonin and dopamine, has been shown to be associated with aggressive behavior [39]; in genetic studies, a low-activity genotype has been associated with higher aggressive behavior [40, 41]. GABA is involved in appraising aggression-related cues, as in step (a) of the social information processing theory in the previous section. Low GABA plasma levels have been found in people with high aggression scores [42]. Lastly, the different neurotransmitter systems work together and with other systems to inhibit or initiate aggressive behavior, and circulating lipids such as cholesterol apparently assist in the communication between the systems. Thus, there is evidence that low levels of different lipids have been associated with aggressive behavior [39].

The next main category of biomolecules involved in aggressive behavior are hormones: sex steroid hormones, stress hormones and insulin. Regarding testosterone, this is the primary steroid hormone responsible for growth and maturation of the male (and to a lesser extent female) body, and has been positively associated with aggressive behavior, in animals more so than humans [43]. Although much attention has been given to this hormone [44],

meta-analysis has only shown a weak positive relationship between testosterone and human aggression [43, 45]. More recently, a study has shown evidence of testosterone being positively linked to not only antisocial, but also prosocial, behavior [46], suggesting a much more complicated role than originally suggested. Regarding stress hormones, the most important one in humans is cortisol, and while results are somewhat inconclusive (especially with respect to reactive aggression), generally a pattern of reduced cortisol is associated with increased aggressive behavior [39]. This suggests that some aggressive behavior results from not assessing risk in a normative way, failing to mount a proper stress response to aggressive stimuli. Lastly, insulin can play a role in aggressive behavior, among both diabetics [16, 47] as well as non-diabetics, with insulin / blood glucose levels being associated with measures of anger / anger expression, hositility, and irritability [47-50]. One study with a modest sample size suggested that the association between insulin and hostility / anger expression may be mostly seen among women and not men [50].

Lastly, some cytokine inflammatory biomarkers have been shown to play a role in aggressive behavior. Several studies have shown positive associations of C-reactive protein and interleukins with aggressive behavior [39, 51, 52]. The links between cytokines and mechanisms with aggressive behavior are still under investigation, but one study [51] showed that a sample of chronically physically aggressive males exhibited reduced levels of pro- and anti-inflammatory cytokines; elevated cortisol levels were also found. Cortisol has been shown to regulate inflammatory responses [53]. Additionally, cytokines have been shown to influence serotonin [54].

As seen, there are already some well-researched biological pathways for aggression. However, it seems there is much overlap and communication between these systems and these connection(s) are still being investigated. Furthermore, our earliest and most well-known molecular pathways (the neurotransmitters and testosterone) have actually shown a weak ability to explain the variance in aggressive behavior, including regarding sex differences. For example, a meta-analysis of serotonin–aggression associations showed only a weak negative correlation ($r=-0.12$), noting that effect sizes have diminished over time and that discrepancies exist in the direction or strength of the association based on the rater (self-ratings vs other raters)[55]. Thus, there is still much to be done regarding the biology of aggression. While it is important to advance our understanding of the already known biomarkers of aggression, new ones also need to be identified to continue to unravel the mysteries of this complex human behavior.

2.1.5. Development and Stability of Aggressive Behavior

One thing that is important to keep in mind about aggressive behavior is that nearly everyone behaves aggressively at some point(s) in life. The behavior itself is not necessarily pathological, there are some normative levels of aggressive behavior in certain situations and over the human lifetime. It is when people display levels higher than their peers, and for long periods of time, that it is labeled it as abnormal, maladaptive, and/or psychopathological.

Longitudinal studies have identified that the peak in physical aggressive behavior is in very young age – 18–36 months – and then generally tapers off as we get older [1, 56]. Developmental psychologist Richard Tremblay even argues that humans are physically aggressive from birth, but that due to babies being small, weak and not intending to harm, we tend not to attribute infant acts as aggressions, although we could [1]. For indirect aggression, the peak in behavior is in middle childhood [57]. However, as we age, we acquire more self-control over our actions and emotions and can more easily adhere to social norms, thus the general decline in the behavior over time. This, at least, is the normative pattern. There has also been suggestive evidence that over historical time, levels of aggressive behavior have decreased in general [58, 59].

Regarding sex/gender differences, it is consistently shown that males exhibit higher levels of aggressive behavior than females from at least schoolage onward [26, 60], with evidence from younger ages being mixed [61]. Furthermore, Archer [29] has shown that gender differences in physical aggression are largest in early adulthood. Findings have generally shown that males are more likely to employ direct aggression and females indirect aggression [25-28], though a meta-analytic review called the indirect aggression gender difference trivial [7]. Verbal aggression does not appear to show a gender difference [25]. Furthermore, it has been shown that children tend to be consistent (over time) with the type(s) of aggression they use [62]. Regardless, a meta-analysis has suggested that only 5% of the variance in aggression is explained by gender differences [63]. Additionally, there have been suggestions that the noted sex differences are diminishing over time [59, 63], although this interpretation has been cautioned [64]. Complicating our study of sex differences in aggression is that most early studies focused on male-only datasets, however in recent decades there has been a large effort to include females in datasets, and more recently to even investigate female-only datasets.

Studies have also indicated a general rank-order stability in individuals' levels of aggressive behavior [65, 66], so that if someone is in the top 10% of aggressive behavior levels at age 3, they are also likely to be in the top 10% at age 7. Thus, while the overall aggressive behavior levels of a cohort are

decreasing over time, there are certain individuals who are generally performing more of the acts within each age (sometimes described as the continuity–discontinuity paradox). However, there are some caveats to note in this pattern. First, while in adulthood individual aggression levels are very stable, and rank-order stability is high [4], the patterns in childhood are less uniform and robust. Childhood aggressive behavior levels do not correlate well with adult aggressive behavior levels, except among the most aggressive males and children of lower socio-economic status [4, 67, 68]. Patterns of continuity from childhood to adolescence, and (separately) from adolescence to adulthood do appear robust for males, but less so for females [4, 26]. Kokko and Pulkkinen [26] showed that 18% of the variance in adult aggression was explained by aggression at age 8–14 years. Taken together, at least adolescent aggressive behavior appears a good predictor of adulthood aggressive behavior, especially with multiple assessment methods (e.g., using multiple raters of behavior) and multiple timepoints [4].

Additionally, some research has suggested that there is generational stability of aggressive behavior, so that children with parents and grandparents who have higher levels of aggressive behavior often have higher levels of aggressive behavior themselves [69, 70]. Explanations for this include both a social learning perspective and genetic perspective [65]. Indeed, the heritability of aggression in children and adults has been shown to be around 50% [71], indicating a strong genetic component to this behavior though the genetic basis remains to be revealed. Additionally, while sex differences in the genetics of aggressive behavior have been shown [72], a larger longitudinal study concluded there were no qualitative sex differences, though quantitative sex differences in heritability did exist, however, they generally decreased with age [73].

Lastly, during adolescence, some of the abovementioned developmental patterns become complicated. In pioneering psychologist Terrie Moffitt's oft-cited theory [74] on antisocial behavior (of which aggressive behavior can be a part, see Figure 1), she describes the paradox of antisocial behavior showing both high continuity and prevalence changes over age, with a high temporary peak in adolescence. She describes that some antisocial behavior is normative in adolescence, but goes on to delinate two specific problematic types of antisocial behavior. The life-course persistent (or, early-onset) pattern more often leads to pathological problems and negative outcomes, while the adolescence-limited pattern is more temporary. However, focusing on the aggressive behavior piece of antisocial behavior, studies have found no strong evidence of physical aggression onset after age six [60], meaning this adolescence-only pattern is unsupported for physical aggression [65].

2.1.6. Predictors and Outcomes of Aggressive Behavior

While thus far I have described aggressive behavior in general, here I start to focus on problem levels of aggressive behavior (i.e., above normative levels; I will look at how problem levels are measured in section 2.5.1) and look at what comes before (predictors) and what comes after (outcomes) chronically exhibiting aggressive behavior. Although Tremblay et al. [56] has shown that by age 17 months 80% of children's mothers reported onset of physical aggression, the vast majority of young children will desist in their aggressive behavior before school age; only about 2–16% of children will have aggressive behavior problems in childhood [75]. However, the earlier the onset of problem levels of the behavior, the higher the risk for persistent/chronic aggressive behavior [60, 65]. It is important to understand the factors that help us identify those children who will go on to have serious problems related to aggression. These factors can be called predictors or indicators or antecedents or risk factors, but its important to remember that these are not for normative levels of aggressive behavior, these are for levels of aggressive behavior that go on to cause problems for the individuals, their social circle (i.e., family and friends), and their community (e.g., school or neighborhood).

In Table 1, I have gathered together a sample of studies that have identified predictors of problem aggressive behavior. This is not meant to be an exhaustive or systematic review. As mentioned in section 2.1.1, the definitions used to study aggressive behavior can be too narrow or wide, making uniform review of this topic challenging. I here generally tried to gather studies that had performed longitudinal analyses so that they were able to examine a timepoint before aggressive behavior problems became chronic. I also focused the predictor investigation on aggressive behavior or conduct problems only (not psychiatric diagnoses or comorbidities, where possible), so as not to have a definition of aggression that was too broad. Many studies looked at specific subtypes of aggression, and some may have results specific to only one gender (as noted).

Table 1. Predictors of problem aggressive behavior, by subtype

Aggression Subtype	**Predictor** (ref)
Aggression, non-specific	Fearlessness [76]; Stimulation-seeking [76]; Large body size [76]; Change in family structure (e.g., divorce) [77*]; Negative parenting [78*, 79*]; Family stress [78*]; Family criminality [79*]; Alcohol use [80*]; Low family income [79*];

	Poor academic achievement/Low IQ [79*, 81, 82]
Reactive	Exposure to traumas, neglect, isolation [21, 83]; Attention problems [83]; Strong emotionality [83]; Poor executive function [84][a]
Proactive	Aggressive role models [21]
Direct, Physical	Male gender [57, 67]; Poor executive function [84][b]; Low maternal education [67, 85*]; Maternal antisocial behavior [86]; Mother who had children at young age [57, 86]; Maternal smoking during pregnancy [86]; Poor parenting strategies (hostile, coercive, ineffective) [57, 67, 86]; Family dysfunction [86]; Low family income [67, 86]; Having a younger sibling [86]
Indirect	Poor executive function [84]; Female gender [57] *For girls*: Low family income [27]; Low parental social support [27] *For boys*: Parenting issues (poor interaction, inconsistency) [27]

*Male-only sample or male-focused review

[a]Study also tested proactive aggression, but did not find the association

[b]Connection to physical aggression was mediated through habitual anger experience

From Table 1, most predictors of problem aggressive behavior are related to the family (e.g., home environment, maternal factors, parenting styles), personal traumas, and personality/temperament traits. Additionally, gender differences, if able to be measured at all (some studies were male-only datasets), appeared mixed. For example, one study [84] found no gender differences, while other studies appeared to find gender differences for physical and indirect aggression. Gender differences were not noted regarding reactive or proactive aggression.

In understanding the predictors of problem aggressive behavior, who to target for preventive and intervention measures can be clarified to try to avoid future harm. This harm comes in many forms. Of course, there is the harm that comes to the victim, but aggressive behavior is also a problem for the aggressor and society. Table 2 presents a selection of the outcomes related to aggressive behavior. This is again not meant to be an exhaustive or systematic review. I

here gathered studies that had performed longitudinal analyses on aggressive behavior or conduct problems. Some studies looked at specific subtypes of aggression, and most looked at datasets with both genders (as noted).

Table 2. Outcomes of aggressive behavior, by subtype

Aggression Subtype	**Outcomes** (refs)
Aggression, non-specific	Peer rejection [87]; Strained child–teacher relationship [87]; School maladjustment [88]; Educational attainment [89]; Long-term unemployment [88]; Divorce [90]; Substance (ab)use [3, 88, 91-95]; Risky sexual behavior [3, 94]; Youth pregnancy [3]; Mental health problems [3, 91, 96*]; Attempted suicide [91]; Traffic violations [69]; Delinquency[a] [95, 97]; Crime [3, 91, 96*, 98, 99]; Domestic violence [3, 69]
Reactive	Peer rejection [83]
Proactive	Peer rejection [83]; Delinquency[b] [100*]; Crime [101][b]
Direct, Physical	*For boys*: Criminal offending [60]

*Male-only dataset

[a]For Roff et al., the association is only for males

[b]The studies also specifically found *no* significant link between reactive aggression and delinquency/crime

From Table 2, outcomes of aggressive behavior involve many areas of life, from increased risky behavior (substance use and sexual health risks) which can lead to substance abuse and teen pregnancy, to mental health problems/disorders, to criminal offending or violence against oneself and others. Additionally, within the listed studies, gender differences appear mixed. For example, some studies [3, 95] noted no gender differences in outcomes, while others [60, 97] noted that delinquency/criminal offending was more common for highly aggressive males than females. Lastly, regarding the entertwined aggression–IQ/academic performance association, while negative education-related outcomes are seen on Table 2 (and the related outcome of unemployment), Fergusson et al. [3] showed that once adjusted by confounders (i.e., adverse childhood environment/events, demographics, co-occuring behavioral/emotional problems, and cognitive

ability), the adult education/employment outcomes of childhood conduct problems became statistically non-significant.

On a positive note, Kokko et al. [88] have shown that child-centered parenting and prosocial behavior can lower the risk of aggressive children becoming long-term unemployed in adulthood.

2.2. Behaviors and Problems Related to Aggression

In this section, I widen the scope of research to situate aggressive behavior in the greater context of human behavior. In 1966, psychologist Thomas Achenbach coined the terms 'externalizing behaviors' (of which aggressive behavior is a part; Figure 1) and 'internalizing problems', which are used to characterize many of the child/adolescent social adjustment and mental health problems [102]. I will focus on these categorizations, and also discuss prosocial behavior, which is a positive and beneficial behavior relating to social adjustment and mental health. Additionally, I discuss how these are related to each other generally, and to aggressive behavior specifically.

2.2.1. Externalizing Behaviors

In Achenbach's original analysis [102], using child psychiatric case reports, he amassed a list of symptoms and separated them into the broad-band dichotomous categories 'externalizing' and 'internalizing' by using factor analysis. Since the categorization was first introduced, more than 75,000 publications have utilized it [103]. Externalizing symptoms are disinhibited externally focused behaviors, and can be aggregated into aggressive behavior and rule-breaking behavior and sometimes hyperactivity and attention problems. The categorization is hierarchical in that you can investigate specific symptoms, or aggregate symptoms into narrow-band subcategories such as 'aggressive behavior' or aggregate all externalizing symptoms together into the broad-band 'externalizing' category. The reasons to aggregate or not depends on the research question being asked, since there are etiological and other distinctions between the (sub)categories [103-106]. Additionally, an emerging concept aggregating both externalizing and internalizing categories under one 'general psychopathology' dimension – the p factor – will also be discussed in section 2.3.1.

To illustrate the subcategories of externalizing behavior, I provide an example question from Achenbach's well-known and well-used Child Behavior Checklist (CBCL) [107] that gets at each (other questionnaires will be discussed in section 2.5.1). For aggressive behavior, a typical question asks if the child 'threatens people'; for rule-breaking behavior, if the child engages in

'lying or cheating'; for hyperactivity, if the child 'talks too much'; and for attention problems, if the child 'can't concentrate/can't pay attention for long'. A review has shown that boys tend to score higher on externalizing problems than girls, across multiple raters and multiple countries [108].

Because of how the externalizing category was created and conceptualized, it is perhaps not surprising that externalizing behaviors are correlated with each other. One study [109], for example, found correlations ranging 0.48 to 0.76 between aggressive behavior and rule-breaking behavior, depending on sex and rater, and another [110] reported correlations of 0.54 between aggressive behavior and rule-breaking behavior and 0.55 between aggressive behavior and attention problems, using the CBCL. Direct/physical aggression, in particular, has been shown to co-occur with other externalizing problems [7, 111]. Additionally, Pulkkinen [23] showed that proactively aggressing boys were prone to other externalizing problems.

Regarding outcomes, children with aggressive behavior and other co-occurring externalizing problems often have more negative outcomes, such as peer rejection, school drop-out, unemployment, delinquency, and substance (ab)use [112-114], which are often more persistent or severe [85]. One female-only study [111] found that aggressive–hyperactive schoolgirls were at increased risk for later teen pregnancy and being on welfare assistance. Additionally, a study [115] showed that aggressive behavior with hyperactivity was associated with impaired cognition, while aggressive behavior only was mostly situation-specific and did not show association with impaired cognition.

Although aggressive behavior alone does occur, co-occurrence with other externalizing behaviors seems common (10–20% vs 80–90%, respectively [116]). Thus far, research has generally focused on the more severe and psychiatric comorbid problems and not as much on patterns of co-occurrence with general externalizing behavior, which can be helpful to provide context for the more serious problems. Additionally, research on gender differences related to co-occuring behaviors is less common.

2.2.2. Internalizing Problems

Internalizing symptoms are overinhibited or internally focused problems, and can be aggregated into anxiety, depression, and somatic problems. Somatic problems include symptoms like headaches, stomach aches, and sleep problems. Anxiety can be characterized in many different ways, in broad (e.g., including depressive symptoms: anxious/depressed) or narrow (e.g., social anxiety) categories. To illustrate the subcategories of internalizing behavior, I provide an example question from the CBCL [107] that gets at each. For anxiousness, a typical question asks if the child 'fears he/she might think or

do something bad'; for depressiveness, if the child is 'withdrawn or doesn't get involved with others'; and for somatic problems, if the child experiences 'aches and pains without medical reasons'. A review has shown that girls tend to score higher on internalizing problems than boys, across multiple raters and multiple countries [108].

Regarding co-occurrence, one study [110] reported internalizing problem correlations of 0.45 between withdrawn/depressed and anxious/depressed, 0.24 between withdrawn/depressed and somatic problems, and 0.37 between anxious/depressed and somatic problems. Furthermore, correlations between aggressive behavior and withdrawn/depressed, anxious/depressed and somatic problems were 0.28, 0.46, and 0.26, respectively [110]. A meta-analytic review [7] showed that indirect aggression, in particular, is associated with internalizing problems. Additionally, Pulkkinen [23] showed that proactively aggressive girls were more prone to internalizing problems.

Regarding outcomes, children with aggressive behavior and co-occurring internalizing problems often have more negative outcomes. One Finnish study [96] showed that boys with both conduct and internalizing problems at age eight were at the highest risk for psychiatric disorders, criminal offenses and self-reported problems in adulthood. Ladd et al. [87] followed young schoolchildren for three years and found that children with both aggressive and withdrawn behavior were most likely to be lonely, dissatisfied, friendless, disliked, victimized, and have a poor student–teacher relationship.

These problems with co-occurrence are important to better understand. Some children with behavioral/emotional problems do seem to have isolated/singular problems, while many have co-occuring internalizing (or externalizing) problems, and there is another group: those who have co-occuring internalizing and externalizing problems (28–57% in [116]). One study [117] showed that 19–23% of their sample of young schoolchildren had only internalizing problems, 21–22% had only externalizing problems, while 48% had both internalizing and externalizing problems. Futhermore, they showed that those with only internalizing problems were the most likely to improve over time, while the kids with only externalizing problems had a 25% probability of moving into the 'comorbid' category. Thus far, research has generally focused on the more severe and psychiatric comorbid problems and, additionally, less research exists on gender differences related to co-occuring behaviors.

2.2.3. Prosocial Behavior

Prosocial behavior does not have any one definition in the literature. It generally encompasses abilities of constructive behavior such as conflict resolution skills, helping, sharing, cooperating and self-control of negative

emotions [4], and is stable over time [118]. A predictor of prosocial skills is a close parent–child relationship [118]. In 1984, Eron et al. suggested that prosocial and antisocial behavior were opposite ends of a single dimension of behavior, since they are consistently negatively correlated [118], however, this has been called into question [119]. Indeed, it has been shown that individuals with aggression problems but without deficits in prosocial behavior exist, and often they have better outcomes [88, 120, 121]. Additionally, low prosocial skills alone have been shown to be predictive of later arrests [118, 120]. In combination, however, individuals with both aggressive behavior and low prosocial skills are the most at risk for later criminality [121].

In general, girls have been shown to have higher prosocial skills than boys [122]. This somewhat relates to the meta-analytic finding that direct aggression (which boys are more likely than girls to exhibit) is associated with poor peer relationships [7]. Additionally, the same meta-analysis also showed that indirect aggression is associated with higher prosocial behavior, while also highlighting that the gender difference in indirect aggression levels is neglible. Thus, the relationship between aggressive and prosocial behavior regarding gender is still yet to be fully clarified.

2.3. Neuropsychiatric Aggression Disorders

Only a fraction of aggressive behavior problems become severe enough and long-lasting enough to require psychiatric assessment/intervention. These severe disorders of aggressive behavior are not the main focus of the thesis work, however, it is important to understand some of these diagnoses in the context of aggressive behavior development, how severe aggressive behavior can get, and that these disorders are generally associated with the more negative outcomes mentioned in section 2.1.6.

2.3.1. P Factor

As mentioned in section 2.2.1, the concept of a 'p factor' was proposed by Caspi et al. [123], aggregating externalizing, internalizing, and thought problem categories under one 'general psychopathology' dimension. This p factor is generally used in the context of mental health disorders, and appears to better explain psychiatric disorders than separating them into the externalizing, internalizing, and thought problem subcategories. The Diagnostic and Statistical Manual of Mental Disorders (DSM) [75], now in its 5th edition, has traditionally listed psychiatric disorders as distinct entities, although in the 5th edition comorbidity was taken into consideration. Indeed, as noted in sections 2.2.1 and 2.2.2, co-occurrence at non-pathological levels between the different behavioral and emotional problems exists, and

Achenbach et al.'s [103] review suggests that analyses should include both internalizing and externalizing problems to be able to control for co-occurrence. Additionally, recent genetic assessment, using four different methods, has shown that the p factor explains 22–57% of the variance in the psychiatric disorders tested (e.g., major depressive disorder, bipolar disorder, substance abuse disorder, attention-deficit/hyperactivity disorder), depending on the method used [124]. However, there is still much variance left unexplained and there are known etiological differences between certain disorders, which still warrants investigating these problems in deaggregated ways alongside aggregation [125].

2.3.2. Adolescent Neuropsychiatric Disorders

In this section, I will briefly review a selection of the most relevent adolescent psychiatric disorders that pertain to aggressive behavior, namely, oppositional defiant disorder (ODD), conduct disorder (CD), attention-deficit/hyperactivity disorder (ADHD) and briefly discuss psychopathic traits (as they pertain to CD).

ODD is defined, in the DSM 5, as a 'pattern of angry/irritable mood, argumentative/defiant behavior, or vindictiveness' lasting at least six months exhibited by at least four of the eight listed behaviors, including 'often argues with authority figures/adults' and 'often loses temper', during interactions that do not exclusively include a sibling [75]. In a nation-wide American survey [2], ODD lifetime prevalence was 11.3% for girls and 13.9% for boys, with prevalence at age 13–14 being 12.0%, age 15–16 12.6%, and age 17–18 13.6%. In a review of ODD including American, Canadian and New Zealand datasets [126], slightly lower prevalences were generally noted and gender differences were either non-evident or slightly more likely in boys. In a review on ODD [127], a debate appears to exist with ODD either being considered somewhat distinct and generally having a good prognosis, and ODD being, in some cases, a stepping stone to CD due to ODD children being four times more likely to go on to have CD than non-ODD children. However, this 'progression' from ODD to CD is less clear for girls [127].

CD is defined in the DSM 5 as a 'repetitive and persistent pattern of behavior in which the basic rights of others or major age-appropriate societal norms or rules are violated' exhibited by at least 3 of the 15 listed behaviors, including 'often initiates physical fights' and 'is often truant from school', during the last 12 months, with at least 1 behavior in the last 6 months [75]. In nation-wide American survey [2], CD lifetime prevalence was 5.8% for girls and 7.9% for boys, with a prevalence of 4.4% at age 13–14, 7.5% at age 15–16, and 9.6% at age 17–18. It also indicated that the median age of onset for behavioral disorders was 11 years. However, one study [128] showed that for children with CD, as age increased, aggressive behavior became less common while

non-aggressive symptoms increased. A review [126] suggested that once ODD or CD behavior is apparent in girls, the stability appears higher than boys. Girls with CD appear possibly more at risk than boys to develop co-occuring internalizing problems [126]. Additionally, Loeber et al. [9, 126] noted that the onset of CD is particularly early in boys with ADHD.

Within the DSM 5 diagnosis for CD, there is a specifier regarding psychopathic traits. These traits include lack of empathy, emotion, or feelings of remorse or guilt (also called 'callous-unemotional' traits) [75]. Blair et al. [129] notes that less than half of children with CD exhibit callous-unemotional traits and often have a worse prognosis than children without these traits (see also [130]). Blair et al. [129] also characterizes CD into two subgroups: those with callous-unemotional traits, and those with heightened threat sensitivity and reactive aggression.

ADHD is defined in the DSM 5 as a 'persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development' as exhibited by six or more inattention symptoms and/or hyperactivity and impulsivity symptoms for six or more months to a degree that is non-normative [75]. In a nation-wide American survey [2], ADHD lifetime prevalence was 4.2% for girls and 13.0% for boys, with a prevalence of 8.8% at age 13–14, 8.6% at age 15–16, and 9.0% at age 17–18. Comorbidity among children with ADHD is common [131], with no significant differences between genders regarding comorbid externalizing disorders [132]. Blair et al.'s [129] review of CD indicated that half of children with ADHD have an anxiety disorder, and more than that have either CD or ODD (see also [133]). The review also discussed that while some of the literature shows a typical progression from ADHD to ODD and then CD (see also [134]), this is not the only pathway to CD and later serious outcomes, nor does it indicate that most children with ADHD or ODD go on to have CD.

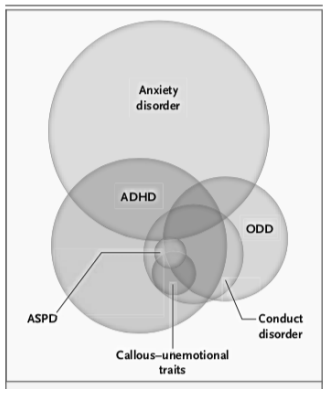
2.3.3. Adult Neuropsychiatric Disorders

In the previous section, psychiatric disorders that are diagnosed in childhood/adolescence were the focus, although ADHD can also be a disorder of adulthood. Since it is often comorbid with CD and ODD, I chose to introduce it in the previous section. I also note that there is an aggressive behavior disorder called intermittent explosive disorder (IED), which can be diagnosed in childhood as well as adulthood. IED involves 'recurrent behavioral outburst[s] representing a failure to control aggressive impulses' as manifested by either verbal or physical aggression (that does not involve damage or destruction) toward property, animals, or people, roughly twice or more per week for a 3-month period, or 3 behavioral outbursts involving injury or destruction to property, animals, or people within a 12-month period [75]. This disorder, however, is not a focus of this thesis. Here, I will mainly

focus on the adult psychiatric disorder called antisocial personality disorder (ASPD) and its relationship with aggressive behavior.

ASPD is one of ten personality disorders classified in the DSM, and is part of the B cluster (of A, B, and C) that includes borderline, narcissistic, and histrionic disorders. The DSM 5 says ASPD is a 'pervasive pattern of disregard for and violation of the rights of others', with evidence of conduct problems typically before age 15, as indicated by 3 or more of 7 listed behaviors (e.g., irritability and aggressiveness, deceitfulness, impulsivity, consistent irresponsibility). The individual has to be at least 18 years old to receive the ASPD diagnosis and the occurrence of behavior cannot happen only during bipolar or schizophrenic episodes. Approximately 50% of children with disruptive disorders (e.g., CD, ADHD) go on to have ASPD (see Figure 3, for a visual representation of the associations between ASPD and CD, ODD, ADHD) [129, 133]. In an American review of the epidemiology of ASPD [135], while 25% of adults exhibit 'syndromal antisocial behavior' only 4.3% of the adult population has ASPD: 6.4% for men, 2.4% for women. Among the prison population, however, 47% of men and 21% of women have been found to have ASPD [136]. A later age of onset for ASPD for women is indicated in the literature [137]. Individuals with ASPD are at increased risk of disability, excess mortality, medical illnesses and injury, as well as criminality [135]. ASPD responds poorly to treatment, however, the symptoms typically improve or cease by midlife [135], although this interpretation is cautioned by Balsis et al. [138] who suggest that there may be measurement bias in the DSM criteria, as it mainly contains young adult behaviors.

Figure 3. Diagram of overlap among six clinical entities (anxiety disorder, callous-unemotional traits, conduct disorder, ADHD, ASPD, ODD)



ADHD: attention-deficit/hyperactivity disorder; ASPD: antisocial personality disorder; ODD: oppositional defiant disorder

Size of each circle represents the relative prevalence of the clinical entity in the community. Reproduced with permission from *New England Journal of Medicine* [129], Copyright Massachusetts Medical Society.

Comorbidity of ASPD with other disorders (such as substance use disorder, generalized anxiety disorder, depressive disorder, and posttraumatic stress disorder) is also common [135]. As with CD, there is a substantial portion of individuals with ASPD who have psychopathic (callous-unemotional) traits (Figure 3)[129]. Because of the co-occurrence with other disorders and symptoms, some have suggested that there may be subtypes ASPD [22], but using latent factor analysis, Bucholz et al. [139] showed that findings for both men and women do not support this, instead the disorder appears to exist along a continuum of severity [140].

Regarding prediction of ASPD, the picture is incomplete. Studies often investigate only antisocial behavior (not the ASPD diagnosis) or have used male-only, prison population or clinical population samples, which hampers generalizability. Gender differences in ASPD require clarification because although women are less likely to have ASPD, a Finnish prison study on homicide offenders found that ASPD increased the odds of being an offender by 10-fold in men, but 50-fold in women [141]. Additionally, predictors of ASPD still require clarification. Caspi et al. [91] has shown that undercontrolled (i.e., externalizing problems) 3-year-olds have a small but increased risk of ASPD later in life. Some studies have shown that CD but not ADHD predicts ASPD [130, 142, 143], while some have shown that both predict ASPD [81]. Depressive symptoms [77, 130] and callous-unemotional traits [130] have also been shown to predict ASPD. Finally, regarding subtypes, both reactive and proactive subtypes have been associated with ASPD [144], and covert but not overt CD symptoms improve prediction of ASPD [142].

2.4. Assessment of Aggressive Behavior

As seen in the above sections, there are many different aspects to consider when studying aggressive behavior. Thus, it is likely unsurprising that there are numerous ways to assess aggressive behavior. Here, I will discuss the relevant measurement instruments (questionnaires and interviews) and biological measurement approaches, as well as discuss differences that arise based on our choice of who rates the aggressive behavior and in which population it is studied in. It is important to understand the tools to capture the different aspects of aggressive behavior so that studies can be properly designed to most effectively improve our understanding of aggression.

2.4.1. Aggressive Behavior Instruments

There are dozens if not 100+ instruments to capture and characterize aggressive behavior. One review on aggressive behavior assessment reminds

us that some measurement methods capture trait aggression, others state aggression, and that there are different methods of data collection: self-report, observer-reported, interviews, projective tests, and laboratory methods [145]. Another review focusing on proactive and reactive aggression presents some instruments that exist to capture those subtypes specifically [33]. Some instruments capture related concepts like anger, hostility, psychopathology, antisocial behavior, and are used to study aggression. There is also the purpose of the questionnaire to consider; is it for tracking psychological development and maladaptive deviations, a screening tool for psychological/psychiatric referral, or a proxy for diagnosis?

There are six different instruments for aggressive behavior assessment used in this thesis, many of them are well-known and commonly used for psychological development, psychopathological, and/or psychiatric assessment purposes. In this section, I will briefly introduce each one.

The Achenbach System of Empirical-Based Assessment (ASEBA) was designed to measure behavioral and emotional problems in childhood, adolescence and adulthood and comprises a family of questionnaires for different raters at different ages of the individual in question [146, 147]. It is one of the most well-known and well-used instruments available for psychopathological research, used in over 57 countries [108, 148]. Achenbach (the system's creator) is often reviewing the system and the literature and making new recommendations and improvements, usually for clinical purposes [103, 148-151]. This thesis utilizes the Child Behavior CheckList (CBCL), the Youth Self-Report (YSR), the Teacher Report Form (TRF), and the Adult Self-Report (ASR). These questionnaires are approximately 100 items in length, and based on factor analysis, approximately 8 syndrome scales and 2 broadband scales are formulated (varies depending on the questionnaire). The broadband scales (Externalizing and Internalizing) encompass some of the syndrome subscales (Externalizing: Aggressive Behavior (~20 items) and Rule-Breaking (~12 items); Internalizing: Anxious/Depressed (~16 items), Withdrawn/Depressed (~8 items), Somatic Complaints (~9 items)), with the remaining syndrome scales being Attention Problems (~26 items), Social Problems (~11 items), and Thought Problems (~1 item). Subscale score cut-offs based on gender and country exist for suggesting a child is in the 'normal', 'borderline', or 'clinical' range.

The Autism–Tics, ADHD and other Comorbidities inventory (A-TAC) captures neurodevelopmental disorders and associated conditions in children [152]. It can be used as a diagnostic proxy (with existing cut-offs), but is not a stand-alone diagnostic tool. The A-TAC is a comprehensive screening interview via the telephone that includes 96 items comprising 20 modules, which correspond to diagnostic domains such as Conduct Disorder (5 items),

Attention Deficit/Hyperactivity Disorder (19 items), and Oppositional Defiant Disorder (5 items).

The Devereux Child Behavior (DCB) captures problem behavior in children [153]. It is used to assess psychopathology based on DSM criteria. The short version of the DCB includes 42 items comprising 7 scales: Emotional Liability (5 items), Social Isolation (3 items), Aggressive Behavior (7 items), Attention Problems (5 items), Anxiety Problems (6 items), Dependency (5 items), and Physical Coordination (5 items). This questionnaire has evolved since its creation into the Devereux Behavior Rating Scale and the Devereux Scales of Mental Disorders.

The Multi-dimensional Peer Nomination Inventory (MPNI) was developed by Finnish developmental psychologist Lea Pulkkinen for tracking longitudinal psychological development [104]. It has not been used outside of Finland, though it has been used in cross-cultural comparisons with other instruments [89, 121, 154]. Compared to the CBCL, the MPNI covers a more broad range of child social adjustment (both active and passive) [104]. The MPNI is a 37-item questionnaire originally developed for use among peers of school-age children, but has been modified for other raters: parents, teachers, and children about themselves or their siblings (see Appendix 1). The subscales of the MPNI were derived from factor analysis, which identified three main factors: behavioral problems (i.e., externalizing problems), emotional problems (i.e, internalizing problems), and adjustment (i.e., prosocial behavior). These factors can be further subdivided into 9 subscales: aggression (6 items), impulsivity-hyperactivity (7 items), inattention (4 items), depression (5 items), social anxiety (2 items), victimization (1 item), constructiveness (5 items), compliance (4 items), social activity (3 items). Additionally, the aggression subscale comprises both direct (4 items) and indirect (2 items) subtypes of aggression.

The Strengths and Difficulties Questionnaire (SDQ) was designed to be a short (25 item), user-friendly tool to measure psychosocial problems and strengths in children and youth [155]. The SDQ comprises five scales with five items each: Conduct Problems, Hyperactivity-Inattention, Emotional Problems, Peer Problems, and Prosocial Behavior. The externalizing subscale comprises Conduct Problems and Hyperactivity-Inattention, and the internalizing subscale Emotional Problems and Peer Problems. The SDQ creator (Goodman) indicates that for many purposes the SDQ functions as well as the ASEBA system, correlating highly with it, but is shorter to administer [156]; it can be used as a screening tool.

Lastly, the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) is a one-on-one psychiatric interview [157, 158]. It is a validated multi-diagnostic instrument that collects information on several psychiatric

disorders based on DSM-IV criteria. Although the instrument most extensively covers alcohol-related disorders, it also covers other psychiatric disorders such as CD and ASPD.

Many of these instruments are typically used for clinical screening or proxy diagnosing, but others function to describe a wider range of normative to maladaptive levels of child behavior and problems (e.g., ASEBA, MPNI, SDQ). Some studies exist to directly compare [149, 159] or even harmonize [160] the more well-known instruments, but since so many instruments exist, this effort is challenging and has been rare. Achenbach et al. [161] suggests that cross-instrument (cross-rater) correlations are moderate (r=0.3) for externalizing and internalizing problems. One area where this comparison data becomes important is in phenotyping in biomarker and (epi)genetic studies. In order to obtain quality biomarker or genetic results, a clear understanding of the phenotype being studied is required since most biomarker or genetic studies require pooling together large amounts of participants in order to find meaningful results, and the samples often include studies that used different instruments. Without knowing how the instruments capturing 'aggressive behavior' are related to each other, it makes interpretation difficult.

2.4.2. Approaches to Discover Biomarkers of Aggression

Just as there are multiple instruments for collecting phenotypic aggressive behavior data, there are many ways to investigate aggression from a biological standpoint. Typical methods have been candidate biomolecular pathways and genes, or molecular experiments using animal models [162, 163]. In recent decades, (epi)genome-wide association studies have been used [71, 164], and more recently biomarkers have been a focus [39]. The term biomarker can be used very broadly (e.g., to include epigenetic marks), but here it refers to metabolites, the result of the use of the proteins that are coded for by DNA/RNA in biological pathways. Metabolites are intermediate or end-products in pathways that can provide information on not only what genes are involved in a phenotype, but which part(s) of the biological pathway are affected. Biomarkers have been the focus of research for many diseases, and have more recently been used in psychiatry [165].

The two main approaches for characterizing biomarkers involve either mass spectrometry (MS) or nuclear magnetic resonance spectroscopy (NMR). These two techniques are complementary, with differing strengths and limitations (see Table 3). NMR has been gaining in popularity in biomarker research during recent years due to its ease of sample handling, ability to use a variety of fluids such as blood, urine, cerebrospinal fluid, and lower time and monetary costs [166, 167]. It also allows for a hypothesis-free approach to investigating biological systems, something that is useful when trying to generate new hypotheses that can then be followed up more deeply with

molecular or other types of approaches. A 2016 review suggested that NMR should be applied to metabolite investigations pertaining to aggression [39], however few such studies have thus far been published [168].

Table 3. Comparison of mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR)[a]

Approach	Strength	Limitation
MS	More sensitive than NMR	Possible bias from chromatographic separations
	Covers wider range of metabolites than NMR	Uncertainty in metabolite ability to create a detectable peak
	High selectivity (well-suited to targeted analyses)	Demanding sample preparation and measurement
NMR	Minimal and easy sample handling	Often 10–100 times less sensitive than MS
	Easy quantitative analysis of metabolite concentrations	Non-selective analysis (peaks overlap)
	Built-in redundancies for metabolite identification	
	High level of reproducibility	
	Non-destructive to sample	

[a] based on information from [166, 167]

2.4.3. Assessments by Different Raters

To obtain information on aggressive behavior, it needs to be decided who to collect data from. The individual in question can self-report on their behavior, but depending on their age/maturity, personal bias, and societal stigma possibly leading to underreporting their problem behaviors, they may not be the most reliable source [5, 169, 170]. Thus, information on aggressive behavior has also been gathered from parents, teachers, siblings, partners, peers, mental health professionals and impartial observers. There is also the consideration to collect information from a person of the same or opposite gender to the individual in question [169]. It appears that certain observers have been used more or less frequently, depending on the field or type of study. For example, epidemiology studies seem to typically use parent, self and, to some extent, teacher ratings. Informant choice can be due to many

factors; for example, it is often easier to obtain self and parent ratings, and more time-consuming and cumbersome to obtain teacher or peer ratings.

Studies have suggested, however, that collecting data from multiple raters is ideal [150, 171-173]. One reason is that situational aggression means that individuals are not always aggressive in all environments of their life, although findings suggest that the more settings an individual is aggressive in the more generalized and thus more serious their problem usually is [172]. Another reason is that different raters have different perspectives; for example, a teacher regularly sees the wide range of normative and non-normative behavior levels of children at a similar age/maturity level, while peers can see children in settings away from adult observers where they may behave differently. Thus, different observers might actually report different information from each other, providing a more holistic picture of the behavior.

Although it is recommended to collect data from multiple raters, not all studies do so, and those that do sometimes ask: Which rater(s) are 'best' to use or best able to predict X outcome? And, how should using multiple raters be handled? This is a whole line of research in itself [174], but I will briefly outline the factors that pertain to aggressive behavior regarding multiple raters. First, because (many forms of) aggressive behavior are overt, informants/raters are more likely to agree on this behavior compared to other behavioral and emotional problems [5, 104, 161, 171, 175, 176]. Studies have generally found that the highest rater agreement (r range roughly 0.50–0.75) occurs between similar raters: mother and father, a teacher and another teacher, and self (test–retest) [5, 109, 177]. Additionally, teacher and peer ratings have moderately high correlations (approximate r range of 0.50–0.60)[104, 176, 178]. Rater combinations such as parent/teacher and parent/self and teacher/self are generally the weakest, with correlations approximately in the 0.20–0.30 range [5, 175, 176], and children typically rating higher aggression levels than adult observers [173, 177] and teachers indicating lower levels of aggressive behavior than other informants [66, 179].

Regarding gender differences, some studies have found no or minimal differences in male and female ratings by informant [5, 104, 171], while others have found significant interaction between gender and rater [29, 169, 173, 178]. Rate agreement appears better when the child in question is younger compared to older [5, 173], though findings on this are mixed [171]. Additionally, rater agreement appears to be different whether the sample of children are from a community or clinical setting [175, 180].

In considering informants' utility, some studies have found certain informants to be more useful than others ('useful' usually regarding significant prediction of future outcomes). Some studies have cautioned against using only self reports [5, 169], although they appear to become more useful the older the

individual in question is [161, 169, 176]. Although used less often, teacher ratings are often shown to be among the most useful ratings of aggressive behavior [77, 175, 177, 181-184]. However, teachers seem to rate boys as more aggressive than girls, compared to other informants [171, 182, 185].

Finally, in considering how raters compare with other behavioral and emotional problems, a somewhat different pattern occurs. Internalizing problems, for example, show lower rater agreement than externalizing problems [5, 161, 177], at least in certain sample populations [175]. One study [173] showed a gender-by-informant (and sometimes age as well) interaction with internalizing problems. Regarding prosocial behavior, the data appear mixed, some show moderate correlations between raters [104] while others show low correlations [186].

2.4.4. Assessments in Different Populations

The final area of consideration regarding assessment of aggressive behavior concerns populations. Who should be studied in order to better understand aggressive behavior? Highly aggressive individuals only or everyone? Only males or females too? Adults or children, or both? In this section, I briefly describe some aspects to consider in the assessment of aggressive behavior based on the population studied.

Early aggression research focused more on those who were in the prison system or clinical/psychiatric patients. These clinical or prison populations are important for understanding who the highly aggressive individuals are, what comorbidities they have, what kind of outcomes they experience, and other such important aspects. However, focusing only on these populations, does not help us understand the normative levels of aggression or what the patterns of aggressive behavior are for people with higher levels of aggression and co-occurrence that do not reach clinical levels. School-based, community-based, and population-based samples have been used to this end, with population-based samples often being a gold standard sampling design in epidemiology. Most of the samples in this thesis are population-based.

Furthermore, in population-based sampling, special cases of this sampling design can be considered. The typical population-based design tries to include as many individuals in a defined population (e.g., all adolescents in a country) as possible; to narrow the 'population' down, this often becomes a cohort design in which all people born in a certain year or range of years are targeted for participation. Studies can also selectively include or exclude singletons (individuals born as the only child of a pregnancy) or multiples (e.g., twins, triplets) in order to be able to take relatedness into account in analysis. For example, if you have a mix of related and unrelated individuals in a study, results may be biased by related individuals, unless adjusted for. Some

datasets, therefore, specifically aim to include singletons (or treat individuals as unrelated, by adjusting), while others exclusively enroll twins (or other multiples). There is a long history in the field of behavior genetics of utilizing twin cohorts in a classic twin design [187] for genetically informed analysis, and many of the cohorts used in this thesis involve datasets collected for these purposes. Twins are born, on average, prematurely and at lower birthweight than singletons, however, investigations with cohorts in this thesis have shown that the twins in childhood and adolescence appear similar to singletons regarding behavioral and emotional problems [188, 189]. An additional benefit of twin samples is the generally high motivation to participate, thus high response rates.

The study sampling design is also a factor to consider regarding study populations. If individuals are sampled only once, or if multiple time points/ages are looked at but different individuals in those time points are used, they are being studying cross-sectionally. This study design can be used to look at associations and co-occurrence, but cannot declare anything about cause and effect or prediction. In contrast, longitudinal datasets follow the same individuals over two or more periods of time. This design can be useful to investigate, for example, the development of aggression, factors that occur before the problematic aggression pattern that affect its progression, or factors that contribute to worse outcomes.

The gender of the chosen population is also a factor in aggressive behavior research. Much of the early research on aggression came from male-only datasets, this made sense because of the seemingly high rates of aggression and violence among males, compared to females. However, because of the more nuanced understanding of gender differences mentioned in above sections, it is equally important to understand aggression in females, and across both genders in general. Due to the previous lack of female-only research, there are now investigations pointedly studying aggression among females (e.g., FemNAT-CD consortium; http://www.femnat-cd.eu/), however, it is also necessary to use samples that combine both males and females, studied in the same way at the same time, to be able to better suss out the gender similarities and differences seen in study findings. Thus, population-based samples are well-suited to this.

Regarding age for the study population, this depends on the research question. Developmental aspects of aggressive behavior require young children (or even pregnant mothers), schoolage children, and/or adolescents. Young and older adults can be used in longitudinal studies to study the life-course of aggressive behavior, or in, for example, generational aggression and assortative mating regarding aggressive behavior. This thesis uses children/adolescents in all studies, but also includes young adults in studies on long-term outcomes and biomarkers.

Lastly, in considering issues of generalizability, investigations that are focused in only one country (or region) or that can be done across countries also need to be considered. Achenbach has done lots of work with his ASEBA instruments to assess the system's ability to characterize the same phenotype in different populations [108, 148]. However, most other study instruments have not been investigated as widely. While aggressive behavior is a universal human behavior, it is important to understand what patterns of problems and outcomes hold (or become unclear) across countries, cultures, and school systems.

AIMS OF THE STUDY

The overarching aim of the thesis is to improve our understanding of aggressive behavior in childhood and adolescence by looking at the following aspects: co-occurring behaviors/problems, differences between raters, aggression subtypes, adolescent aggression as a predictor of a young adult personality disorder, and biomarkers of aggression.

Specific thesis aims are:

1) to characterize childhood/adolescent aggressive behavior and its associations with co-occurring behaviors/problems using parental, self, and teacher ratings across genders as well as multiple countries and ages (Study I and II, as well as unpublished supplementary results);

2) to identify if adolescent aggressive behavior (general, and direct and indirect subtypes), independent of other externalizing behaviors, can predict antisocial personality disorder in young adulthood (Study III, as well as unpublished supplementary results);

3) to identify new, potential blood biomarkers of aggressive behavior (Study IV, as well as unpublished supplementary results)

MATERIALS AND METHODS

The aims and studies for this thesis were a part of and grew out of work being done in the ACTION consortium (Aggression in Children: unravelling gene–environment interplay to inform Treatment and InterventiON strategies; http://www.action-euproject.eu/) [190]. ACTION brought together 11 (mainly European) partners with strong data and expertise in genetics, epigenetics, and metabolomics and access to large childhood cohorts gathered as prospective twin studies, population-based studies, and/or clinical studies. The focus of ACTION was to improve our understanding of the causes of individual differences in childhood aggression and co-occurring problems, especially through the use of genetics, epigenetics, and metabolomics (Study IV contributed to this). For all these investigations, the phenotyping of aggressive behavior plays a large role in helping to clarify the results by understanding the similarity and differences in the different cohorts, instruments, and ages of the children (Study I, II, and III contributed to this).

4.1 Study Materials

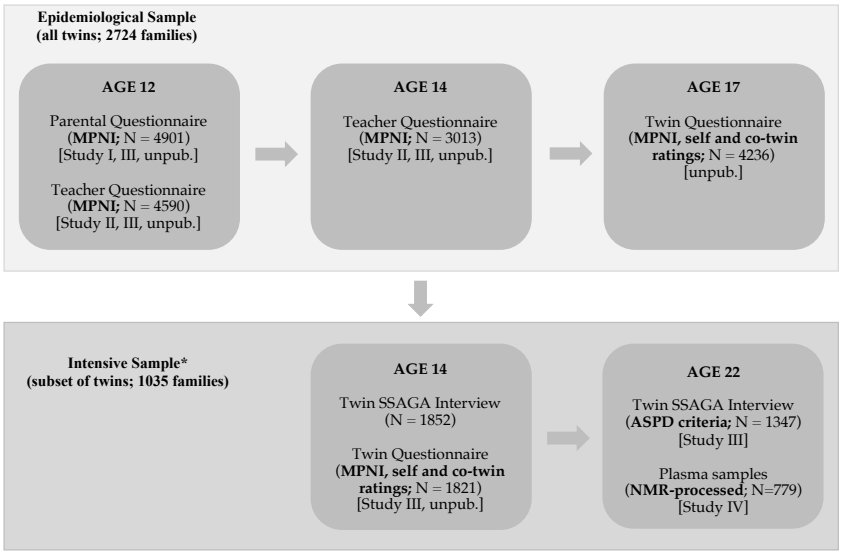
4.1.1 FinnTwin12 (FT12)

The population-based FinnTwin12 (FT12) cohort is a longitudinally collected cohort of Finnish twins born 1983–1987, initiated when twins were age 11/12 years in order to investigate genetic and environmental precursors to health-related behaviors (see Figure 4 for study flowchart) [191, 192]. The Finnish population register was utilized in order to contact all families with twins born in the study years, with 87% responding to participate (2724 families). Data was collected mainly through questionnaires at ages 11/12, 14, 17, and 22 (called the 'Epidemiologic sample'; approximately 5600 twins, with 85–90% response rates).

At age 14, a subset of the twin families (called the 'Intensive sample'; 1035 families) were asked to additionally participate in a more intensive data collection including one-on-one psychiatric interviews using the Semi-Structured Assessment for Genetics of Alcoholism (SSAGA) instrument [157], a saliva sample, and additional questionnaires. At age (approximately) 22, this Intensive group also underwent another SSAGA interview, plasma sample, and additional questionnaires. The plasma sample was processed using the NMR technique described in section 2.5.2.

The FT12 cohort is used in all studies (Study I, II, III, and IV) as well as in the unpublished results, with all analyses based on FT12 data performed by me.

Figure 4. Flowchart of FinnTwin12 study waves and data collection included in the thesis



*Data for intensive sample participants includes both epidemiological sample data collection materials as well as intensive sample data collection materials

Abbreviations: ASPD=Antisocial Personality Disorder; MPNI=Multidimensional Peer Nomination Inventory; NMR=nuclear magnetic resonance spectroscopy; SSAGA=Semi-Structured Assessment for the Genetics of Alcoholism; unpub.=unpublished results

4.1.2. Netherlands Twin Register (NTR)

The Netherlands Twin Register (NTR), established in 1987, longitudinally collects data on twins and other multiples (and their families) from birth onward to examine, in particular, behavioral development and psychopathology [193]. Data collection is nationwide roughly every 2 years from birth to age 18 (at which point they are invited to the adult twin registry). In the adult registry, questionnaires and blood samples are collected in different collection waves. The blood samples are processed using the NMR technique described above in section 2.5.2. In 1999, the NTR began collecting data on the twins from their teachers when twins were ages 7, 10, and 12 years old, with response rates being approximately 60%.

The NTR cohort is used in Study I, II, and IV, with NTR analyses performed by NTR for Study I and by me in Study II and IV (Table 3).

Table 3. Sample sizes for NTR for Study I, II, and IV

	Study I	**Study II**	**Study IV**
Rater-Child Age (Study Instrument)	*Boy N/Girl N*	*Boy N/Girl N*	*Male N/Female N*
Mother-3 (ASEBA)	9277/9360		
Mother-5 (DCB)	7520/7695		
Father-5 (DCB)	6808/6985		
Mother-7 (ASEBA)	5720/5853		
Father-7 (ASEBA)	4134/4182		
Teacher-7 (ASEBA)		3416/3518	
Mother-9 (ASEBA)	4543/4689		
Father-9 (ASEBA)	3210/3255		
Teacher-10 (ASEBA)		3264/3318	
Mother-12 (ASEBA)	3870/4010		
Father-12 (ASEBA)	2764/2839		
Teacher-12 (ASEBA)		2477/2576	
Self - Wave 8 (ASEBA)			551/1260
Self - Wave 10 (ASEBA)			455/1063

Abbreviations: Achenbach System of Empirical-Based Assessment (ASEBA); Devereux Child Behavior rating scale (DCB)

4.1.3. Twins Early Development Study (TEDS)

The population-based Twins Early Development Study (TEDS) cohort is a longitudinally collected cohort of UK twins born 1994–1996, initiated when twins were age 2 years in order to investigate language, cognitive, and behavioral development [194]. Data collection has occurred roughly every 2 years from birth to age 21. TEDS collected data on the twins from their teachers when twins were ages 7, 9, and 12 years old, with response rates being 85%, 76%, and 78%, respectively.

The TEDS cohort is used in Study I and II, with TEDS analyses performed by TEDS for Study I and by me in Study II (Table 4).

Table 4. Sample sizes for TEDS for Study I and II

	Study I	**Study II**
Rater-Child Age (Study Instrument)	*Boy N/Girl N*	*Boy N/Girl N*
Parent-4 (SDQ)	3581/3788	
Parent-7 (SDQ)	2740/2892	
Teacher-7 (SDQ)		2834/2987
Parent-9 (SDQ)	1055/1245	
Self-9 (SDQ)	1055/1245	
Teacher-9 (SDQ)		1295/1467
Parent-12 (SDQ)	1828/2117	

Self-12 (SDQ)	1828/2117	
Teacher-12 (SDQ)		2168/2477
Parent-16 (SDQ)	2134/2632	
Self-16 (SDQ)	2134/2632	

Abbreviation: Strengths and Difficulties Questionnaire (SDQ)

4.1.4. Generation R (GENR)

The population-based Generation R (GENR) cohort is a longitudinally collected cohort of Dutch children born in and around Rotterdam 2002–2006, initiated while the children were still in utero in order to investigate growth, development and health from fetal life to young adulthood [195]. Data collection has occurred during pregnancy, at birth, and frequently throughout childhood. GENR collected data on the children from their teachers at age 7 years, with a response rate of 77%.

The GENR cohort is used in Study I and II, with GENR analyses performed by GENR for both Study I and II (due to ethical permissions restrictions) (Table 5).

Table 5. Sample sizes for GENR for Study I and II

	Study I	**Study II**
Rater-Child Age (Study Instrument)	*Boy N/Girl N*	*Boy N/Girl N*
Mother-3 (ASEBA)	2271/2246	
Father-3 (ASEBA)	1840/1897	
Mother-6 (ASEBA)	2887/2856	
Teacher-7 (ASEBA)		2270/2242
Mother-10 (ASEBA)	2250/2310	
Father-10 (ASEBA)	1624/1670	

Abbreviations: Achenbach System of Empirical-Based Assessment (ASEBA)

4.1.5. Child and Adolescent Twin Study in Sweden (CATSS)

The population-based Child and Adolescent Twin Study in Sweden (CATSS), initiated in 2004, longitudinally collects data on Swedish twins born since July 1992 in order to investigate how both genetic and environmental effects influence health and behavior in children and adolescents [196]. The overall response rate has been 80%.

The CATSS cohort is used in Study I, with CATSS analyses performed by CATSS (Table 6).

Table 6. Sample sizes for CATSS for Study I

	Study I
Rater-Child Age (Study Instrument)	*Boy N/Girl N*
Parent-9 (A-TAC)	5610/5516
Parent-12 (A-TAC)	1649/1598
Parent-15 (SDQ)	2083/2199
Self-15 (SDQ)	2258/2806

Abbreviations: Autism–Tics, ADHD and other Comorbidities inventory (A-TAC); Strengths and Difficulties Questionnaire (SDQ)

4.1.6. Twin study of Child and Adolescent Development (TCHAD)

The population-based Twin study of Child and Adolescent Development (TCHAD) in Sweden, initiated in 1994, longitudinally collects data on Swedish twins born in 1986 in order to investigate how genetic and environmental effects contribute to the development of mental health problems over time [197]. Response rates have been between 67 and 91%.

The TCHAD cohort is used in Study I, with TCHAD analyses performed by TCHAD (Table 7).

Table 7. Sample sizes for TCHAD for Study I

	Study I
Rater-Child Age (Study Instrument)	*Boy N/Girl N*
Parent-8 (ASEBA)	552/534
Parent-13 (ASEBA)	535/522
Self-13 (ASEBA)	560/551
Parent-16 (ASEBA)	532/507
Self-16 (ASEBA)	583/606

Abbreviations: Achenbach System of Empirical-Based Assessment (ASEBA)

4.2 Ethical Approvals

Ethical approval for the FT12 dataset comes from the ethical committee of the Helsinki and Uusimaa University Hospital District and the Institutional Review Board at Indiana University, Bloomington, Indiana, USA. Parents provided informed consent for the twins at ages 12 and 14, and the twins themselves provided written informed consent at ages 17 and 22.

All data from other cohorts were collected under protocol(s) approved by the appropriate local ethics committee(s), and analyses were performed in accordance with the ethical standards established in the 1964 Declaration of Helsinki and its later amendments. Data from the NTR and TEDS cohorts

(used in Study II) was officially requested and granted by data administrators and review boards as appropriate.

4.3 Study Instruments

4.3.1. Multi-dimensional Peer Nomination Inventory (MPNI) (Studies I–IV)

For each item in the MPNI (more information about MPNI in the above section 2.5.1), the child in question is rated on a scale from 0 to 3 (0 being 'does not fit the child at all' and 3 being 'fits the child very well'). In the event of an ambivalent response (e.g., a rater circled both 0 and 1), the average of the two values was taken. Aggression mean scores were obtained as follows: total/general aggression by taking the average value across all 6 items, direct aggression by taking the average value across all 4 direct aggression items, and indirect aggression by taking the average value across the 2 indirect aggression items. Other subscale mean scores (e.g., inattention, social anxiety) were obtained in a similar manner.

MPNI data was available for the FT12 cohort and was used in all studies and the unpublished data in this thesis (More information about MPNI in the above section 2.5.1).

4.3.2. Achenbach System of Empirical-Based Assessment (ASEBA) (Studies I, II, IV)

This thesis included the following ASEBA questionnaires: CBCL for ages 1.5–5-year-olds and 6–18-year-olds (Study I), YSR (Study I), TRF (Study II), and ASR (Study IV) (More information about ASEBA in the above section 2.5.1). For the purposes of continuity across study instruments in Studies I and II, the syndrome scale Attention Problems was considered to be part of externalizing problems.

For each item on the questionnaires, the individual in question is rated on a scale from 0 to 2 (0 being 'not true' and 2 being 'very true or often true'). Scores for each item are then summed over the items that make up the individual subscales.

ASEBA data was available for the GENR, NTR, and TCHAD cohorts and was used in Studies I, II, and IV.

4.3.3. Strengths and Difficulties Questionnaire (SDQ) (Studies I–II)

The SDQ comprises five scales with five items each: Conduct Problems, Hyperactivity-Inattention, Emotional Problems, Peer Problems, and Prosocial Behavior (More information about SDQ in the above section 2.5.1). The SDQ recognizes an externalizing subscale as comprising Conduct Problems and Hyperactivity-Inattention, and an internalizing subscale as comprising Emotional Problems and Peer Problems. In Studies I and II, Conduct Problems is used as a proxy for aggressive behavior, although conduct problems are perhaps more severe/serious than aggressive behavior (and contains items that also pertain to rule-breaking behavior). Likewise, in Studies I and II, Emotional Problems is considered as anxiety problems.

For each item, the child in question is rated on a scale from 0 to 2 (0 being 'not true' and 2 being 'certainly true'). Subscale scores are derived as a scaled mean score of the items that make up the individual subscales.

SDQ data was available for the CATSS and TEDS cohorts and was used in Studies I and II.

4.3.4. Devereux Child Behavior (DCB) (Study I)

The short version of the DCB (used in Study I) includes 42 items comprising 7 scales: Emotional Liability (5 items), Social Isolation (3 items), Aggressive Behavior (7 items), Attention Problems (5 items), Anxiety Problems (6 items), Dependency (5 items), and Physical Coordination (5 items) (More information about the DCB in the above section 2.5.1).

For each item, the child in question is rated on a scale from 0 to 4 (0 being 'never' and 4 being 'very frequently').

DCB data was available for the NTR cohort and was used in Study I.

4.3.5. Autism–Tics, ADHD and other Comorbidities inventory (A-TAC) (Study I)

The A-TAC modules used in Study I include: Conduct Disorder (5 items), Attention Deficit/Hyperactivity Disorder (19 items), Autism (17 items), and Oppositional Defiant Disorder (5 items) (More information about A-TAC in the above section 2.5.1).

A-TAC data was available for the CATSS cohort and was used in Study I.

4.3.6. Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) ASPD Criteria (Study III)

In the FT12, the age 14 and 22 waves of data collection for the intensively studied subgroup involved a SSAGA interview (More information about SSAGA in the above section 2.5.1). According to recommendation by our MD, PhD youth psychiatrist co-author (A Raevuori) on Study III, 15 behaviors from the SSAGA's ASPD criteria were considered to be in the ASPD 'diagnosis' assigned in Study III (see Appendix 2 for details on the criteria). These 15 behaviors corresponded to the 7 criteria laid out in the DSM-IV regarding ASPD. If an individual had a history of conduct problems before 15 years of age (assessed at the age 22 SSAGA interview) and endorsed 3+ ASPD behaviors, they were considered to have ASPD.

4.4 Statistical Methods

Detailed methods for statistical analyses can be found within each publication. Here I report the general workflow of analyses that appear throughout the studies. All Studies utilized Stata, versions 11–16 (Stata Corporation, College Station, TX, USA) for analyses and considered statistical significance to be $p<0.05$.

4.4.1. Mean levels (Studies I–III and unpublished results)

Background analyses for studies included reporting on mean levels of aggressive behavior and other subscales. Study I and II had the same initial protocols (means and correlations) in order to provide some continuity and comparability. Study I and II calculated means and standard deviations of the subscales from all available cohorts, study intruments and ages, separated by gender. Study I utilized parent and self ratings, while Study II utilized teacher ratings. Study II additionally calculated effect sizes (Cohen's *d*) and standard errors for all available cohorts, study intruments and ages, separated by gender.

The unpublished results calculated means and effect sizes (Cohen's *d*) for parent vs teacher age 12 aggression ratings (including for direct and indirect aggression subtypes), and for teacher vs self age 14 aggression ratings (including for direct and indirect aggression subtypes). Paired t-tests were conducted to test if means were statistically significantly different ($p<0.05$) between parent and teacher or teacher and self ratings. This was done for genders combined and for genders separated.

Study III calculated the mean MPNI-based aggression scores among both ASPD and non-ASPD individuals. These means were calculated separately for

each rater and gender, as well as also combining the genders. Effect sizes (Cohen's *d*) were also reported, and adjusted Wald tests were used to calculate statistically significant differences in means between ASPD and non-ASPD individuals.

4.4.2. Correlations (Studies I–IV and unpublished results)

In Study I and II, Pearson correlations between aggression subscales and other available subscales within a study instrument were calculated, separately by gender. For Study II, correlations between subscales other than aggression were also presented in Supplemental material.

In the unpublished results, Pearson correlations were calculated between parent and teacher age 12 aggression ratings (including for direct and indirect aggression subtypes), and between teacher and self age 14 aggression ratings (including for direct and indirect aggression subtypes), separately by gender.

In Study III, Pearson correlations were calculated between the total aggression scores for the different raters.

In Study IV, Spearman correlations were calculated between metabolites, and between aggression and metabolites.

4.4.3. Regression modeling (Study II–IV and unpublished results)

Study II utilized linear regression modeling in order to formally test gender differences (including possible gender interaction) and to assess the amount of aggression variation explained by the independent variable. Before models were run, subscales were standardized (mean=0, SD=1) to allow for comparability where possible. Models included aggressive behavior as the dependent and subscale scores as the independent variable(s). Initial models first included only one independent variable, while follow-up models included multiple independent variables (more than one behavioral/emotional subscale) to examine multiple co-occuring problems with aggression.

Study III utilized logistic regression, with the cluster option to adjust for the relatedness of the twins, with ASPD as the outcome variable and aggression ratings (total, direct, and indirect aggression) from different raters as predictor variables, adjusted by age at time of assessment and gender. Before models were run, subscales were standardized (mean=0, SD=1) to allow for comparability where possible. Initial models used only one rating of aggression at a time, while later models included multiple ratings of aggression or included both direct and indirect aggression scores for a single rater. Gender interactions were also tested. For each model, receiver operating

characteristic (ROC) curves were generated, and the area under the curve (AUC) was calculated to indicate the predictive utility of the different aggression ratings in the prediction of ASPD.

Additionally, Study III included several sensitivity analyses to clarify the ASPD prediction. First, the residual effect of aggression with hyperactivity-impulsivity removed on ASPD was examined to clarify that hyperactivity-impulsivity was not the driver for the ASPD prediction. Second, an internalizing problem (social anxiety) was modeled to demonstrate that it did not predict ASPD (i.e., that the prediction abilities were not broad to many psychopathologies). Third, the ASPD definition was modified by changing the amount of criteria needed for ASPD 'diagnosis' to see if there was a change in prediction if the threshold for the outcome was changed. Next, the aggression subtypes with the residual effect of the other subtype removed were modeled to see if any changes in ASPD prediction occured. Lastly, bias from attrition was clarified and shown not to explain the ASPD prediction.

Study IV utilized linear regression, with the cluster option to adjust for the relatedness of the twins, with a metabolite as the dependent variable (because it was measured at age 22) and aggression as the independent variable (measured at younger ages), adjusted by age, sex, and body mass index (BMI). Before models were run, metabolites were rank-transformed, as that was generally the best solution. In this initial modeling step, only one metabolite appeared to show statistically significant association with aggression, 3-hydroxybutyrate, so it was followed-up on further. In follow-up regression modeling, multiple raters of aggressive behavior were used (separately and combined) and models were further adjusted for leisure-time physical activity (METs), smoking status, alcohol consumption frequency, and self-rated general health. Another follow-up analysis (unpublished) involved looking at direct and indirect aggressive behavior subtypes separately. Additionally, sex interactions were tested for. Lastly, to investigate the robustness of the biomarker association, an NTR dataset with young adult 3-hydroxybutyrate values and aggression ratings was used for replication modeling, with the same adjustments as the FT12 data. Analyses for both FT12 and NTR cohorts were also run using Bayesian modeling.

RESULTS

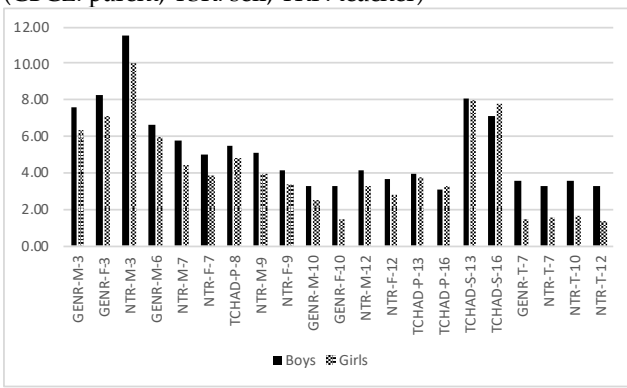
Detailed results for the Studies can be found within each publication. Here I report the main general findings, as well as combine/compare across Studies, where possible. Also, additional unpublished analyses in FT12 are presented.

5.1 Mean levels

5.1.1. Mean levels of aggressive behavior across instruments/cohorts and raters (Studies I and II, and supplementary unpublished analyses)

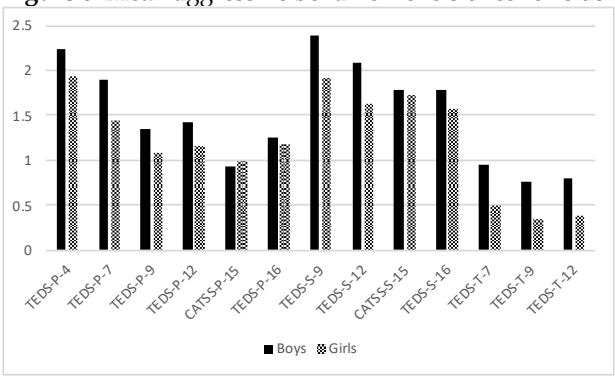
Studies I and II reported mean levels of aggressive behavior as well as many other behavioral/emotional subscales as available in the various study instruments by cohorts. Since this thesis is focused on aggressive behavior, and to reduce the amount of redundant documentation, I here present only the aggressive behavior mean levels, by study instrument and gender (further broken down by cohort, rater, and age). Figures 5, 6, and 7 show the patterns of mean levels of aggressive behavior for the main study instruments (ASEBA, SDQ, MPNI). They are organized to show first all parental ratings (by age), then self ratings (by age), then teacher ratings (by age), in effect showing results from Study I and II together (and unpublished results in Figure 7).

Figure 5. Mean aggressive behavior levels of cohorts using ASEBA study instruments (CBCL: parent, YSR: self, TRF: teacher)



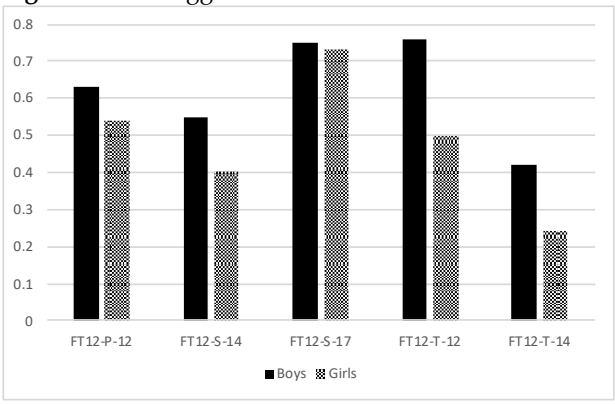
Note: X-axis labels indicate the cohort first, then rater, then age of the child
Y-axis is the mean score on the corresponding ASEBA instrument
Abbreviations: M=mother, F=father, P=parent, S=self, T=teacher

Figure 6. Mean aggressive behavior levels of cohorts using the SDQ



Note: X-axis labels indicate the cohort first, then rater, then age of the child
Y-axis is the mean score on the SDQ instrument
Abbreviations: P=parent, S=self, T=teacher

Figure 7. Mean aggressive behavior levels* of the cohort using the MPNI



*The self ratings at 14 and 17 are unpublished results
Note: X-axis labels indicate the cohort first, then rater, then age of the child
Y-axis is the mean score on the MPNI instrument
Abbreviations: P=parent, S=self, T=teacher

Notable general patterns exist across instruments in Figures 5, 6, and 7, including the general reduction in aggressive behavior by age according to parental ratings, though for self and teacher ratings this pattern is perhaps less clear across study instruments. Across raters, teacher ratings are generally lower than parental and self ratings at similar ages, while parental ratings are notably lower than self ratings at similar ages. Regarding gender differences, the pattern of boys having higher levels than girls emerges, however, this difference is more pronounced in teacher ratings (effect size ranges: 0.34–0.43) and sometimes unclear among self ratings. Nearly all of the gender differences are statistically significant ($p < 0.05$), except for many of the anxiety scales

(FT12 age 14, NTR age 10 and 12, TEDS age 7 and 12) and the somatic scale of NTR age 12.

Notable differences include that while in the same country with the same instrument, ratings for the NTR cohort are generally higher than the GENR cohort for similar ages. Additionally, self ratings among teenagers appear to be more gender equal than younger self ratings across cohorts and study instruments. Finally, FT12 appears to have a reversed pattern regarding parents and teachers; at age 12, teacher-rated aggressive behavior levels are higher than parent-rated levels.

5.1.2. Testing the difference between raters using FT12 (Unpublished results)

Regarding the unpublished results directly comparing FT12 mean levels of parent and teacher aggression ratings at age 12 and teacher and self ratings at age 14, I present Tables 8 and 9, respectively. The tables also compare direct and indirect aggressive subtypes.

Table 8. Mean levels and effect sizes of aggressive behavior for parent and teacher ratings at age 12, by gender and genders combined

Total aggression	**P12 Mean (SD)**	**T12 Mean (SD)**	**p-value[a]**	**Effect size[b]**
Combined	0.59 (0.41)	0.63 (0.65)	0.009	0.05
Boys	0.64 (0.42)	0.76 (0.70)	<0.001	0.18
Girls	0.54 (0.39)	0.48 (0.56)	<0.001	-0.13
Direct aggression	**P12 Mean (SD)**	**T12 Mean (SD)**	**p-value[a]**	**Effect size[b]**
Combined	0.68 (0.48)	0.65 (0.71)	<0.001	-0.08
Boys	0.77 (0.50)	0.86 (0.78)	<0.001	0.11
Girls	0.60 (0.44)	0.43 (0.56)	<0.001	-0.35
Indirect aggression	**P12 Mean (SD)**	**T12 Mean (SD)**	**p-value[a]**	**Effect size[b]**
Combined	0.39 (0.47)	0.59 (0.68)	<0.001	0.32
Boys	0.37 (0.46)	0.58 (0.68)	<0.001	0.35
Girls	0.42 (0.49)	0.60 (0.68)	<0.001	0.29

[a]Paired samples t-test used

[b]Cohen's *d* (positive values indicate teacher rating larger than parent)

Abbreviations: P=parent, SD=standard deviation, T=teacher

Table 9. Mean levels and effect sizes of aggressive behavior for teacher and self ratings at age 14, by gender and genders combined

Total aggression	**T14 Mean (SD)**	**S14 Mean (SD)**	**p-value[a]**	**Effect size[b]**
Combined	0.33 (0.48)	0.48 (0.37)	<0.001	0.31
Boys	0.42 (0.55)	0.56 (0.39)	<0.001	0.23
Girls	0.24 (0.39)	0.40 (0.33)	<0.001	0.43
Direct aggression	**T14 Mean (SD)**	**S14 Mean (SD)**	**p-value[a]**	**Effect size[b]**
Combined	0.34 (0.54)	0.53 (0.42)	<0.001	0.35
Boys	0.48 (0.63)	0.63 (0.44)	<0.001	0.23
Girls	0.22 (0.40)	0.43 (0.37)	<0.001	0.55
Indirect aggression	**T14 Mean (SD)**	**S14 Mean (SD)**	**p-value[a]**	**Effect size[b]**
Combined	0.30 (0.51)	0.38 (0.49)	0.001	0.13
Boys	0.31 (0.52)	0.43 (0.52)	0.003	0.17
Girls	0.29 (0.50)	0.33 (0.45)	0.139	0.09

[a]Paired samples t-test used

[b]Cohen's *d* (positive values indicate self rating larger than teacher)

Abbreviations: P=parent, SD=standard deviation, T=teacher

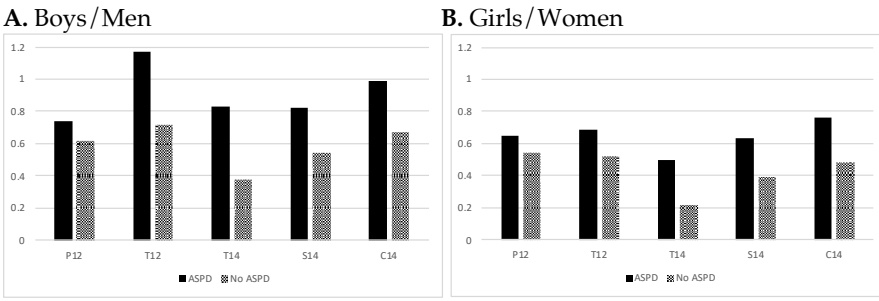
For FT12 parent and teacher age 12 rating mean comparisons, it is important to separate the data by gender due to differences in rater aggression levels and effect sizes. For direct (and total) aggression, aggressive behavior ratings from teachers were higher for boys (parent ratings were higher for girls however). For both boys and girls, teacher ratings for indirect aggression were higher than parents. All parent vs teacher mean differences were statistically significant and effect sizes were all small. Also notable, regardless of rater, direct aggressive behavior is much higher in boys compared to girls, while indirect aggressive behavior levels are more similar between genders.

For FT12 teacher and self age 14 rating mean comparisons, it is again useful to separate the data by gender. Self ratings of aggressive behavior were higher than teachers for all analyses, though girls' indirect aggressive behavior levels were not significantly different between the two raters. Comparing genders, for direct aggression, the effect size for the rater difference was moderate for girls, while it was small for boys. On the other hand, for indirect aggression, the effect size for the rater difference was larger for boys than girls.

5.1.3. Mean levels of aggressive behavior according to ASPD status (Study III)

Regarding mean levels pertaining to aggression and ASPD (Study III), I present a graphical form of table data published in Study III (Figure 8A-B).

Figure 8A-B. Mean levels of aggressive behavior among ASPD and non-ASPD individuals, by gender and rater



Note: X-axis labels indicate the rater first, then age of the child
Y-axis is the mean score on the MPNI instrument
Abbreviations: P=parent, T=teacher, S=self, C=co-twin (sibling)

Here the pattern that boys have higher levels of aggression than girls across raters and ASPD status emerges again. Additionally, higher aggression levels among ASPD individuals compared to non-ASPD individuals are seen (statistically significant p-values <0.01 for boys for all raters except parents, but only statistically significant for girls for self and co-twin ratings, p-values <0.05). Effect sizes between ASPD and non-ASPD individuals (for genders combined) were small (d = 0.32) for parent ratings, moderate for teacher ratings at age 12 and co-twin at age 14 (d: 0.61 and 0.69, respectively), and large for self and teacher ratings at age 14 (d: 0.80 and 0.94, respectively).

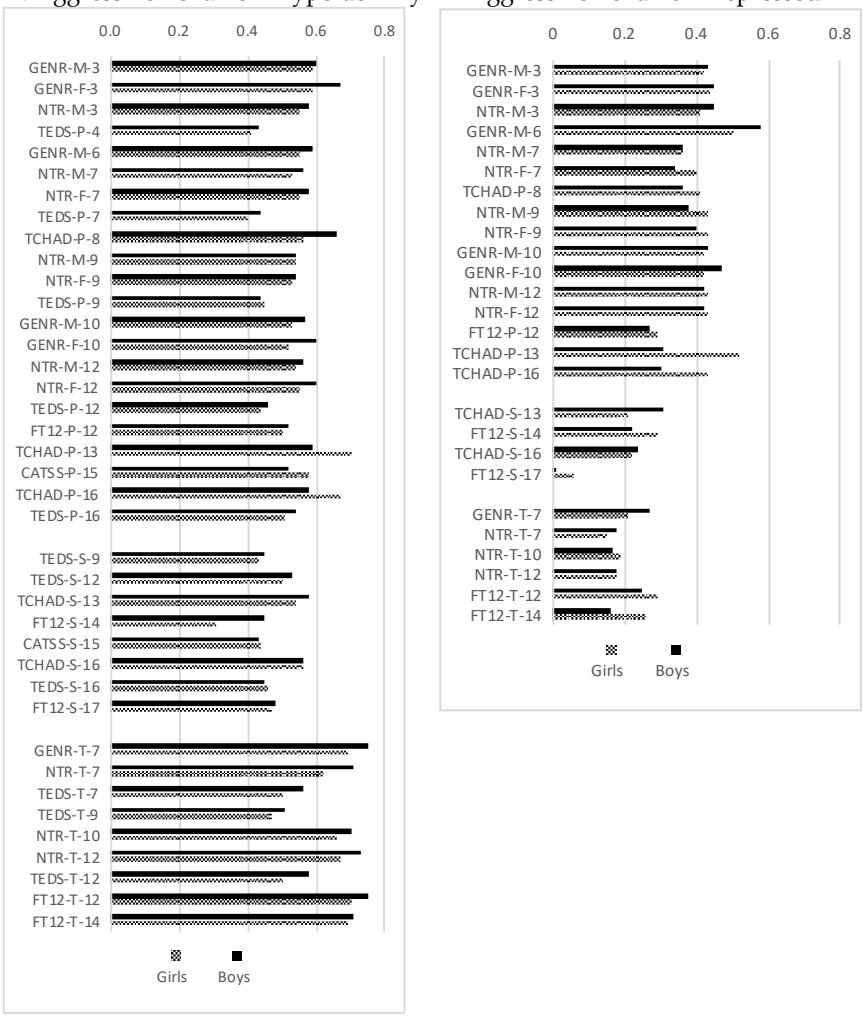
5.2 Correlations

5.2.1. Correlations between aggressive behavior and other behavior/emotional problems across instruments/cohorts and raters (Studies I and II, and supplementary unpublished analyses)

All correlation values for Study I and II can be found in the publications. Here I combine some correlation results from Studies I and II (and supplementary unpublished analyses) and present the correlations of aggressive behavior by selected co-occuring behaviors (Figure 9A-C).

Figure 9A-C. Correlations of aggressive behavior with selected co-occurring behaviors, by rater, age, and gender (includes unpublished FT12 self-ratings)[all significant at $p<0.05$]

A. Aggressive Behavior–Hyperactivity* B. Aggressive Behavior–Depressed^



*'attention problems' from ASEBA instruments as proxy for hyperactivity

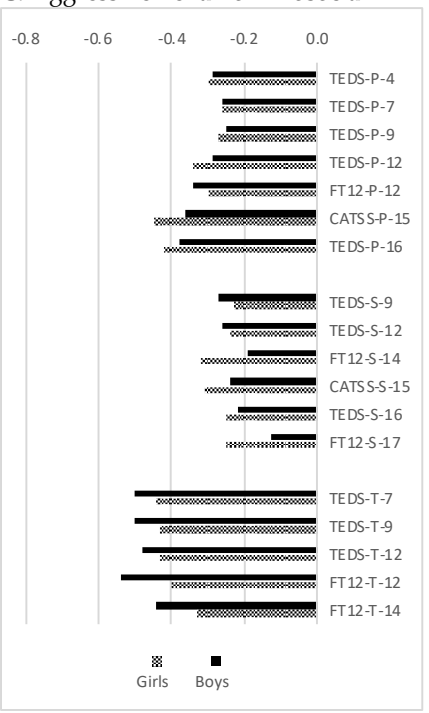
^combines 'withdrawn-depressed' of ASEBA with 'depressive behavior' of MPNI

Note: Y-axis labels indicate the cohort first, then rater, then age of the child

X-axis is the correlation r-value

Abbreviations: M=mother, F=father, P=parent, S=self, T=teacher

C. Aggressive Behavior–Prosocial*



Note: Y-axis labels indicate the cohort first, then rater, then age of the child
X-axis is the correlation r-value
Abbreviations: M=mother, F=father, P=parent, S=self, T=teacher

Across all these datasets and behaviors, some notable similarities emerge. Correlations of aggressive behavior and co-occurring behaviors are generally similar across genders. Across age, within a rater category, correlations are generally similar. For aggressive behavior–hyperactivity (two externalizing behaviors) and aggressive behavior–prosocial behavior, teacher correlations are generally higher than either parent or self ratings. For aggressive behavior–depressed (one externalizing and one internalizing behavior), parent correlations are generally higher than either teacher or self ratings. Correlations among self ratings are generally the lowest, compared to parent and teacher correlations.

There are also some notable differences. For aggressive behavior–hyperactivity correlations, the TEDS cohort generally had the lowest correlations compared to similar aged cohorts. For aggressive behavior–hyperactivity correlations, boys have higher correlations among teacher ratings, while gender levels are more equal for parent and self ratings. For aggressive behavior–prosocial behavior correlations, boys have higher correlations among teacher ratings, while gender levels are more equal for parent ratings and girls tend to have higher correlations among self ratings.

5.2.2. Correlations of aggressive behavior between raters in FT12 (Unpublished results)

Here I present tables showing the correlations between one rater and another for aggressive behavior at age 12 (parent v teacher) and 14 (teacher v self) (Table 10 and 11). The genders have been separated due to the mean level results showing that gender differences exist.

Table 10. Correlations of aggressive behavior for parent and teacher ratings at age 12, by aggression subtype and gender

Total aggression	**Corr**	**p-value**
Boys	0.29	<0.01
Girls	0.20	<0.01
Direct aggression		
Boys	0.29	<0.01
Girls	0.19	<0.01
Indirect aggression		
Boys	0.20	<0.01
Girls	0.17	<0.01

Table 11. Correlations of aggressive behavior for teacher and self ratings at age 14, by aggression subtype and gender

Total aggression	**Corr**	**p-value**
Boys	0.16	<0.01
Girls	0.07	0.09
Direct aggression		
Boys	0.18	<0.01
Girls	0.13	<0.01
Indirect aggression		
Boys	0.05	0.23
Girls	-0.002	0.97

Comparing parent and teacher (age 12) ratings, the correlations for boys were higher than girls for total and direct aggression. For the indirect aggression specifically, correlations were more similar across gender. All correlations were statistically significant.

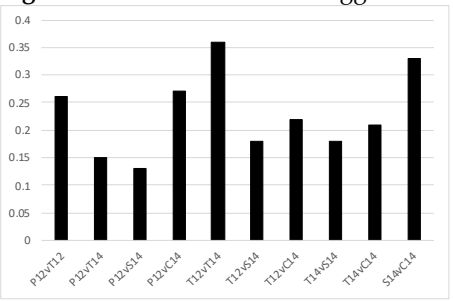
Comparing teacher and self (age 14) ratings, the correlations for boys were higher than girls for direct (and total) aggression scores. Indirect aggression correlations were more similar across gender. Correlations for direct aggression (both genders) and total aggression for boys were statistically

significant, while indirect aggression correlations (both genders) and total aggression for girls were statistically non-significant.

5.2.3. Correlations of aggressive behavior across raters (Study III and unpublished results)

Here I present a graphical representation of correlations of aggressive behavior (total aggression) for different raters from Study III (Figure 9).

Figure 9. Correlations of total aggression between raters*



Note: X-axis labels indicate the two raters being compared; rater first, then age

Y-axis is the correlation r-value

*All correlations significant at p<0.001

Abbreviations: P=parent, T=teacher, S=self, C=co-twin (sibling)

Correlations between raters from Study III are all small to moderate. The largest are between teachers (age 12 v age 14) and between siblings (self v co-twin).

Additionally, Study III calculated the correlations between direct and indirect aggression subtypes (Table 12).

Table 12. Correlations between direct and indirect aggressive behavior, genders combined and separated (separated gender data is unpublished)

Rater	Correlation (gender combined)	Correlation (boys)	Correlation (girls)
P12	0.39	0.42	0.41
T12	0.62	0.65	0.69
T14	0.61	0.66	0.62
S14	0.40	0.40	0.38
C14	0.56	0.60	0.54

*All correlations significant at p<0.001

Abbreviations: P=parent, T=teacher, S=self, C=co-twin (sibling)

5.2.4. Correlations between metabolites and between metabolites and aggressive behavior (Study IV)

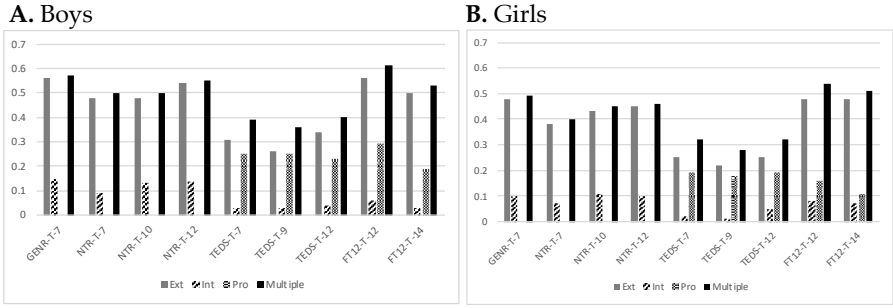
The correlations of the metabolites with each other are well-known associations based on their biological pathways. The metabolite correlations for the branched-chain amino acids were high (r range: 0.62–0.73) as well as for ketone bodies 3-hydroxybutyrate and acetoacetate (r=0.70), but other correlations were weak to moderate (-0.38–0.55). For aggression and the metabolites, correlations were weak, with the highest values being aggression with isoleucine (r=0.14), 3-hydroxybutyrate (r=-0.13) and tyrosine (r=0.13). These three highest correlations were all significant at $p<0.01$.

5.3 Regression Modeling

5.3.1. Proportion of variation in aggressive behavior explained by modeled factors (Study II)

In Study II, data on the initial models with aggressive behavior as the dependent variable and another behavior/problem as the independent variable are presented separately from the data on models with aggressive behavior as the dependent and multiple behaviors/problems as simultaneous independent variables. Here, I present some ways to view the changes that occur when going from models with a single independent variable to multiple independent variables (Figures 10A-B and 11A-B).

Figure 10A-B. R^2 values for the models with only one independent variable (externalizing, internalizing or prosocial behavior) and multiple independent variables, by gender

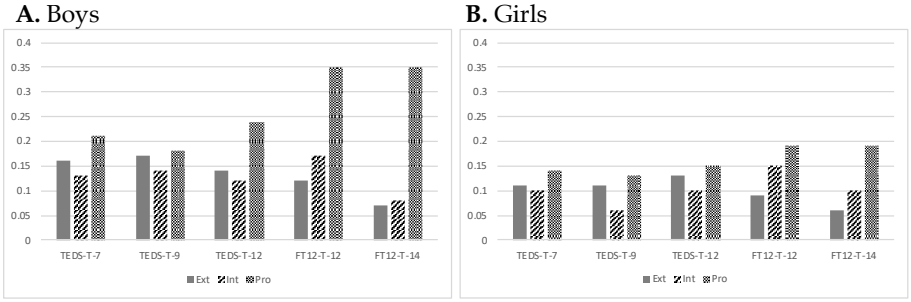


Note: X-axis labels indicate the cohort first, then rater, then age

Y-axis is the R^2 values

Abbreviations: Ext=externalizing (hyperactivity for TEDS and FT12, attention problems for GENR and NTR), Int=internalizing (anxious/depressed for GENR and NTR, depressive behavior for FT12, anxiety for TEDS), Pro=prosocial, T=teacher

Figure 11A-B. Reduction in betas* (initial model value – multiple model value) from the models with only one independent variable (externalizing, internalizing or prosocial behavior) and multiple independent variables, by gender



*All betas were statistically significant in initial and multiple variable models (except age 7 and 9 TEDS betas for anxiety among girls, where betas were significant in initial models and non-significant in the multiple variable model)

Note: X-axis labels indicate the cohort first, then rater, then age

Y-axis is the difference between the initial model's beta value and the multiple model's beta

Abbreviations: Ext=externalizing (hyperactivity for TEDS and FT12), Int=internalizing (depressive behavior for FT12, anxiety for TEDS), Pro=prosocial, T=teacher

By looking at the R^2 values from the different models (Figure 10), the single externalizing behavior as the independent variable explains the largest amount of the variation in aggression, prosocial explains a moderate amount (sometimes nearly as much as the externalizing behavior), while internalizing problems explain little. When multiple independent variables are in the same model, the R^2 value does not increase much above the R^2 value for the externalizing alone model, unless prosocial is in the model, but R^2 increases are still generally small. Additionally, R^2 values are generally higher for boys than girls. Lastly, by looking at changes in effect size (betas), girls had smaller changes than boys when going from initial to multiple variable models, and prosocial behavior betas had the largest changes when going from initial to multiple variable models (Figure 11).

5.3.2. ASPD prediction (Study III)

For Study III, many different main models were analyzed as well as several sensitivity analyses. For complete results, please refer to the publication. Here I present a selected list of progressive and informative models, listing the predictor's odds ratios (OR), 95% confidence intervals (CI), and model AUC values with 95% CIs (Table 13).

Table 13. Select logistic regression models with ASPD as the outcome and main predictors either in a model alone or with multiple predictors included simultaneously*, adjusted for gender and age

MODEL	OR	95% CI	AUC	95% CI
Gender + age	2.1	1.2, 3.4	0.60	0.54, 0.66
T14TA + gender + age	1.7	1.4, 2.1	0.72	0.65, 0.80
T14HI + gender + age	1.7	1.4, 2.1	0.73	0.66, 0.80
T14TA (HI removed) + **gender** + age	1.4 2.2	1.1, 1.8 1.2, 4.0	0.65	0.56, 0.74
S14TA + gender + age	1.8	1.5, 2.2	0.72	0.66, 0.79
S14HI + **gender** + age	1.8 1.9	1.5, 2.3 1.1, 3.1	0.72	0.65, 0.78
S14TA (HI removed) + gender + age	1.6	1.3, 2.0	0.68	0.60, 0.75
T14DA + T14IA + gender + age	1.7	1.2, 2.3	0.73	0.65, 0.81
S14DA + **S14IA** + gender + age	1.5 1.4	1.2, 1.8 1.1, 1.7	0.72	0.66, 0.79
T14TA + **S14TA** + gender + age	1.6 1.9	1.3, 2.0 1.5, 2.4	0.80	0.73, 0.87
P12TA + T12TA + **T14TA** + **S14TA** + C14TA + gender + age	1.5 1.7	1.2, 1.9 1.3, 2.2	0.82	0.76, 0.88
P12TA + T12TA + **T14TA** + **S14TA** + C14TA + P12HI + S14HI + C14HI + gender + age	1.5 1.5	1.1, 2.0 1.1, 2.1	0.82	0.76, 0.89

*Bolded model variables were significant ($p<0.05$) and correspond to the OR value reported in the table, in the case of two significant variables, the OR of the first bolded variable is given first, and the second second.

Note: Model variable labels indicate rater first, then child age, then MPNI subscore

Abbreviations: AUC=area under the curve, C=co-twin (sibling), CI=confidence interval, DA=direct aggression, HI=hyperactivity-impulsivity, IA=indirect aggression, OR=odds ratio, P=parent, S=self, T=teacher, TA=total aggression

From Table 13, in the initial model with only gender and age, gender predicted later ASPD, but the AUC value is low (60%, 95% CI 0.54–0.66). When either

aggressive behavior or hyperactivity-impulsivity are modeled alone, they predicted future ASPD about equally well. Importantly, in models when the effect of hyperactivity-impulsivity on aggressive behavior is removed, aggressive behavior in adolescence still predicted ASPD in young adulthood (though with a lower AUC value). The most parsimonious model included self and teacher aggression ratings at age 14 and predicted ASPD with an AUC value of 0.80 (95% CI 0.73–0.87). Even when hyperactivity-impulsivity predictors were added, they were not significant and the AUC value rose very little, while both self and teacher aggression ratings at age 14 remained significant.

Direct aggression in adolescence predicted future ASPD (for both teacher and self ratings), and indirect aggression (from self ratings at age 14) also predicted ASPD.

5.3.3. Biomarker modeling and replication (Study IV)

For Study IV, the initial regression models looked at all metabolites. The beta coefficients of aggression (metabolites were the dependent variable) were nearly all positive in value and statistically non-significant. However, 3-hydroxybutyrate was statistically significantly (negatively) associated with aggression in linear regression (for detailed Figures and Tables, see the publication). Thus, this metabolite was followed up further.

In FT12 regression models with 3-hydroxybutyrate as the dependent variable and aggressive behavior ratings from different raters as the independent variables (in separate models) adjusted for age, sex, and BMI, the negative association trend holds for nearly all raters (Table 14). Two raters of aggressive behavior showed statistically significant associations with 3-hydroxybutyrate (teacher age 12, self age 14). The percent of 3-hydroxybutyrate value variation explained by aggressive behavior (and age, sex, BMI) was less than 5% in the models, however (conversely, the percent of aggressive behavior variation explained by 3-hydroxybutyrate value, age, sex, and BMI ranges 0.05–9.0%). In models fully adjusting for age, sex, BMI, physical activity level, smoking status, alcohol consumption frequency, and self-rated general health, the negative trend remains (with teacher age 12 and self age 14 still statistically significant; R^2 still less than 5%). Additionally, when putting both teacher age 12 and self age 14 ratings into the same model (since correlation between the two ratings were low), both were statistically significant. Lastly, both direct and indirect aggression appear to be negatively associated with 3-hydroxybutyrate (Figures 12A-B).

Table 14. Standardized beta coefficients and R^2 values[a] for aggressive behavior with biomarker 3-hydroxybutyrate[b]

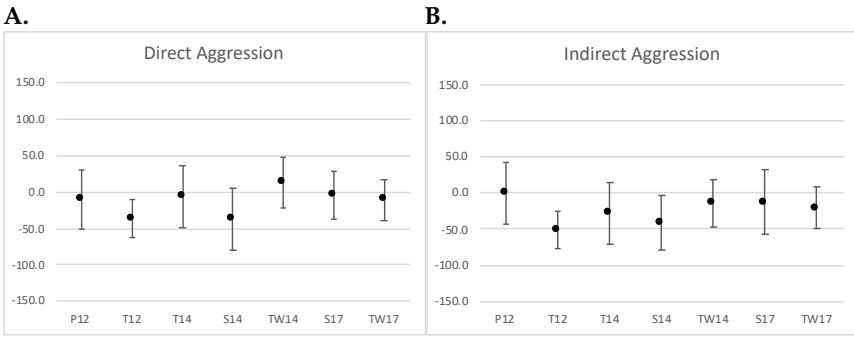
Aggression rating	Standardized Beta Coeff (p-value)	R^2
AggCombined	-0.13 (0.004)	0.040
P12	-0.02 (0.7)	0.022
T12	-0.14 (0.001)	0.039
T14	-0.03 (0.5)	0.022
S14	-0.09 (0.02)	0.029
TW14	0.006 (0.9)	0.021
S17	-0.02 (0.7)	0.018
TW17	-0.04 (0.3)	0.017

Abbreviations: AggCombined=aggression combined; P12=parent rating age 12; S14=self rating age 14; S17=self rating age 17; T12=teacher rating age 12, T14=teacher rating age 14, TW14=co-twin rating age 14, TW17=co-twin rating age 17

[a]R^2 represents the variation explained from all variables in the model together

[b]Table modified from Study IV publication, Supplementary Table 3

Figure 12A-B. Beta coefficients (and 95% confidence intervals) of direct (A) and indirect (B) aggression ratings as (separate) independent variables in linear regression models with rank-transformed 3-hydroxybutyrate as the dependent variable, adjusted for age, sex, and BMI.



Abbreviations: P12=parent rating age 12, T12=teacher rating age 12, T14=teacher rating age 14, S14=self rating age 14, TW14=co-twin rating age 14, S17=self rating age 17, TW17=co-twin rating age 17

In the NTR replication cohort, the same negative association between 3-hydroxybutyrate and aggressive behavior was found. However, statistical significance was lacking in the initial model (adjusted for age, sex, BMI, and batch), but approached significance (p=0.052) among males in the fully adjusted model (age, sex, BMI, batch, physical activity level, smoking status, alcohol consumption frequency, and self-rated general health). Significant sex interaction was noted in the fully-adjusted NTR sample (p=0.015), which was

not found in the FT12 sample (p-value range = 0.10–0.84). Results from Bayesian modeling of both FT12 and NTR models supported the frequentist model results, with models from FT12 now indicating a significant sex interaction: females having a significantly higher negative association with the biomarker compared to males. In NTR, the opposite pattern was found, however, with males having a significantly higher negative association with the biomarker compared to females (who had a small positive association).

DISCUSSION

6.1 Main findings of the thesis

Specifically, this thesis showed that aggressive behavior in childhood co-occurs with other behaviors/emotional problems, with some children having more than one co-occurring behavior. The co-occurrence patterns appear similar across gender, age, and measurement instrument/country, while differences exist between raters. Furthermore, general aggression levels (as well as direct aggression specifically) in pre- and early adolescence can predict ASPD in young adulthood in both genders; when multiple raters of aggression were included, the prediction accuracy improved. While both aggression and hyperactivity/impulsivity were separately able to predict ASPD, aggression without the influence of hyperactivity/impulsivity was still able to predict ASPD, and in models with both aggression and hyperactivity/impulsivity, aggression ratings were the only significant independent variables. Lastly, the ketone body 3-hydroxybutyrate appears to be negatively associated with aggressive behavior; this trend held across raters and in both basic and fully adjusted models, and was supported in a large independent sample and using Bayesian modeling.

6.2 Interpretation of the results and comparison to earlier studies

6.2.1 Characterizing childhood aggressive behavior across genders, subtypes, raters, ages, and countries (Aim 1 + 2)

Looking across gender, aggression subtypes, raters of behavior, age, and countries, results from across the thesis can be compared with each other and fit in with previous studies. Uniformly, results showed the general trend of mean aggressive behavior levels being higher for males than females, which is well-supported in previous studies [26, 60]. However, this finding has several caveats in these thesis findings, as discussed further.

Regarding aggression subtypes, levels of aggressive behavior were different according to gender, rater, and age. For direct aggression the results were consistent, males had higher levels of aggressive behavior than females for all raters and ages, which is consistent with the literature [7, 29, 57]. However, at age 12, among boys, teachers reported more direct aggressive behavior than parents, while among girls, it was the opposite. At age 14, for both boys and girls, self ratings of direct aggressive behavior were higher than teachers. This

may suggest that girls are better able than boys to self-regulate direct aggression at school, but not in other environments. A meta-analysis [7] noted boys as having higher levels of direct aggression than girls for all raters, but (gender difference) correlations were stronger in teacher ratings (r=0.34) than self (r=0.21) and parent ratings (r=0.15). For indirect aggression, Study III showed mixed results with girls at age 12 (ratings by parents and teachers) higher than boys, but at age 14, the levels of indirect aggression were either similar across genders (ratings by teacher and co-twin) or boys had higher (self ratings).

These mixed gender results seem to fit with the meta-analysis by Card et al. [7] showing only trivial differences between the genders regarding indirect aggression. Unpublished results showed teachers (age 12) reporting higher levels of indirect aggression than parents (age 12) for both genders, but self (age 14) ratings were higher than teacher (age 14) ratings (only significant for boys). The school environment may be better for observing indirect aggression than home (where fewer peer/friend interactions occur), although perhaps teachers carry some bias toward characterizing more indirect aggression among girls than boys (and vice versa). Lastly, the Card et al. meta-analysis [7] also indicated an intercorrelation between direct and indirect aggression of 0.82 for girls and 0.88 for boys, which is higher than those found in Study III (and unpublished results): r=0.38–0.42 for parent and self ratings (both genders), r=0.54–0.69 for teacher and co-twin ratings (both genders). These differences could be due to the use of different instruments and differences in raters.

Rater differences between general measures of aggressive behavior were noted, as seen in previous studies. We saw that the highest interrater correlations were between teachers (age 12 v age 14) and between self and co-twins (age 14), while the lowest were between parent (age 12) and age 14 teacher or self ratings. This is in line with the oft-cited meta-analysis by Achenbach et al. [5] that showed that raters who observe children in similar environments have higher interrater correlations. However, results all show lower interrater correlation values than those in Achenbach et al. [5] and other studies [177, 198]. Although, these studies were examining 'externalizing behavior' in general, not aggressive behavior specifically, which fewer studies on interrater correlations have done [109, 176]. Externalizing behavior ratings in general, and aggression specifically, seem to produce the highest interrater correlations, perhaps due to the more easily observable overt nature of many of the behaviors. Thesis analyses showed that age 12 teacher–parent and age 14 teacher–self correlations were highest for boys, perhaps indicating the more overt nature of male aggression or bias in reporting male aggressive behavior. Parents generally reported lower levels of aggressive behavior than teachers (for both genders), while self reports showed higher aggressive behavior levels than teachers (for both genders). Additionally, the greatest gender

differences in aggressive behavior levels were seen in teacher ratings. Taken together, these results support previous suggestions of different raters capturing different levels of behaviors across different environments [172]. Additionally, they show that gender bias may exist in behavior ratings [183, 199], at a rater or societal level, and that the overtness or self-control aspects of behavior development may contribute to gender differences in behavior ratings in different settings [200].

Regarding age trends in aggressive behavior, on average, there is reduction in aggression levels as children get older [1, 56]. This was best observed in Study I's parental ratings. Gender differences were smallest during adolescent ages (compared to younger ages), in particular, self and parent ratings in adolescence were more gender-similar than at younger ages. In a meta-analytic review of sex differences in aggression [29], the first notable sex differences arise in young childhood (as young as 2 years old) for physical aggression, with sex differences in verbal aggression following a couple years later, while the most pronounced sex differences are seen in young adulthood [29]. However, in the teenage years, risky and antisocial behavior is more normative than at other times in development [74]. The highest levels of externalizing problems in adolescence are typically registered from self ratings than from other raters [173, 179]; thus, study results appear to be in line with these known patterns. The minimal gender differences in adolescence highlight that aggressive behavior in girls should not be ignored or downplayed.

Finally, patterns of aggression across country (or, instrument) were generally quite similar. Since each country used separate instruments, it is not possible to specifically indicate whether patterns are similar across countries, instruments, or both. However, similar patterns of childhood behavior across instruments and cultures using ASEBA and SDQ instruments have previously been shown [149], and MPNI items have been compared to other measures in cross-cultural studies, showing comparable behavioral patterns [89, 201]. Thus, our study findings of cross-cultural, cross-instrument similarity are in line with previous findings. Of note, there were some pattern differences between GENR and NTR cohorts (both in the Netherlands, both used ASEBA instruments). However, they could possibly be explained by either the populations the cohorts drew participants from (GENR captures a local population–representative cohort of children in and around the urban area of Rotterdam [only 48% of the target population of children were ethnic Dutch [202]] while the NTR captures children that are representative of the entire national population) or from the type of pregnancy the children experienced in the cohorts (singletons vs multiples). However, studies in FT12 and NTR have shown that twins are not different from other children regarding externalizing/internalizing problems or educational achievement [188, 189],

and patterns were similar between twin cohorts NTR and TCHAD from different countries using the same ASEBA instrument.

6.2.2 Characterizing childhood aggressive behavior and co-occurring behaviors/problems across genders, raters, ages, and countries (Aim 1)

When examining aggressive behavior and co-occurrence with other behaviors or emotional problems, one pattern that emerged was that gender differences were minimal, in teacher ratings in particular. The highest co-occurrence correlations for both genders were for aggressive behavior and other externalizing behaviors, as well as the highest explained variance (R^2) of aggressive behavior. This is well in-line with previous findings [110, 203] and the construct of externalizing behaviors in general [102]. For aggressive behavior co-occurrence with externalizing, and prosocial, behaviors teachers generally had the highest correlations, followed by parents, and then self ratings, for both genders. McConaughy et al. [204] saw minimal externalizing–aggression co-occurrence differences by rater, however, but no formal testing of co-occurrence differences across raters was done here.

Conversely, for co-occurrence of aggressive behavior with internalizing behaviors parents generally had the highest correlations, then self, and teacher ratings, for both genders. This is in line with the internalizing–aggression rater patterns seen in McConaughy et al. [204]. Finally, for both genders, teacher-ratings modeling of aggressive behavior using multiple co-occuring behaviors, internalizing behavior seemed to mainly co-occur with aggressive behavior along with other co-occuring externalizing or prosocial behavior issues. This is in line with other studies [131, 205]. Interestingly, Loeber et al. [126] indicated that the gender with the lowest prevalence of a psychiatric behavioral or mood disorder (e.g., CD being less common in girls) is more at risk to develop a co-morbid condition; so, girls with aggressive behavior may be more at risk for additional psychiatric problems. However, Loeber was looking at DSM diagnoses; Study I and II looked at non-pathological levels of behaviors and emotional problems, which showed similarities in co-occurrence across gender.

Low prosocial behavior, in multiple co-occurrence modeling, was usually associated with aggressive behavior along with other externalizing behaviors, in boys more so than girls. Hämäläinen and Pulkkinen [120, 206] have shown that children with aggressive behavior problems and low prosocial behavior are more likely to be convicted of criminal activities later in life, especially those with multiple problems [121], while aggressive behavior without co-occurrence did not lead to elevated risk of negative outcomes. Kokko and Pulkkinen [88] also showed that aggressive children with low prosocial skills were at increased risk for unemployment later in life, but those aggressive children with prosocial skills were not at increased risk. Thus, aggressive

behavior with co-occuring behaviors is an important consideration for future outcomes.

The patterns of aggressive behavior and co-occurrence appeared similar across age, especially for externalizing and prosocial behaviors. However, these patterns were not tested longitudinally. Previous studies investigating trajectories of co-morbid externalizing behaviors do support the similarity over time, especially among higher levels of aggressive and other problem behaviors [111, 207]. Individuals with continuity in co-occuring problems, being a high-risk group for negative outcomes [111, 133, 207], would benefit from holistic, early interventions.

Lastly, there was also consistency in the results of patterns of aggressive behavior and co-occuring behaviors across countries (or, instruments). Large population-based multi-country (multi-instrument) examinations of these patterns is uncommon in the literature [149], and the similarity of results highlights the robustness of the observations. This signals a possible consideration of studies with single populations/contexts validating their findings in other populations/contexts. Additionally, the robustness indicates that children with multiple co-occuring behavioral issues are common among those who have any behavioral issues, and these children could benefit from interventions that not only address the behavioral and emotional problems, but also provide skills for positive social interaction, such as that from the Fast Track PATHS curriculum [208] or the Strong African American Families (SAAF) program [209]. Finland also has a program (called KiVa) aimed at bullying that targets the entire student body to teach social-emotional skills and provide anti-bullying training [210]. Schools are a good setting to deliver this holistic type of prevention program since school reaches nearly all children in a community, and teachers indicate students of both genders experience these multiple co-occuring issues.

6.2.3 Predicting young adult ASPD using adolescent aggressive behavior (Aim 2)

ASPD prediction showed important considerations regarding gender, aggression subtype, rater and co-occurrence. Two-thirds of the ASPD cases were men, and 6–7% of men had ASPD while only 3% of women. This is in-line with a national epidemiological study of ASPD in the US [135]. Although gender differences in levels of aggressive behavior exist, gender did not play a strong role in prediction. For both genders, aggressive behavior levels in adolescence were higher in participants with young adult ASPD compared to non-ASPD participants. Previous male-only cohorts longitudinally studying ASPD have also shown higher levels of aggressive behavior in childhood [143, 211], but similar studies with females are scarce [3]. In terms of ASPD prediction, when gender and age are in the model alone, gender is a significant

(p-value) factor in prediction, however, the AUC of 0.6 indicates that prediction based on gender is only marginally better than chance (which has an AUC of 0.5). Thus, to predict ASPD better, other factors are needed in the model. ASPD prediction improved when both aggressive behavior and gender were in the model, with gender then only being significant in models with age 12 raters (and no gender interactions were found).

While the main findings of Study III involved the total aggression score, the subscores of direct and indirect aggression subtypes were also utilized. Direct aggressive behavior was shown to consistently and significantly predict ASPD (especially with the age 14 teacher and self ratings). Interestingly, indirect aggressive behavior also indicated possible prediction of ASPD, with self ratings at age 14 being significant in models. The small number of questions that create the MPNI indirect aggression subscale could have led to the weaker trend in prediction, future studies could investigate these subtypes more thoroughly. To my knowledge, ASPD prediction studies involving subtypes of aggression are rare.

As seen from Studies I–II, different raters identify different levels of aggressive behavior, and for Study III some raters contributed more than others to improving ASPD prediction. In prediction models, teacher and self ratings at age 14 were significantly associated with young adult ASPD, with AUCs greater than 0.70. In models where multiple raters were included (because inter-rater correlations were low to moderate), teacher and self ratings at age 14 were both still significant and the AUC rose above 0.80. This model with multiple raters (AUC 0.80, 95% CI 0.73, 0.87) was significantly better at predicting than the parental ratings alone model (AUC 0.65, 95% CI 0.59, 0.72), from the fact that the CIs do not overlap. While parental ratings are used often in psychopathology research, these self and teacher ratings perhaps provide more value in ASPD predictions, and including multiple raters (depending on who they are) can improve models over single ratings.

Finally, Study III investigated whether behaviors other than aggression could predict ASPD or were influencing the ASPD prediction. Results showed that adolescent social anxiety was unable to predict future ASPD, however, hyperactivity/impulsivity could (AUCs greater than 0.70). This provides empirical support for the schematic relationship of anxiety, ASPD, and ADHD in Figure 3. This also aligns with a previous finding that 3-year-olds with externalizing problems were at increased risk for later ASPD [91], and a previous study showing both CD (a proxy for aggressive behavior) and ADHD (a proxy for hyperactivity) predicting ASPD [81]. Because Studies I–II showed that aggressive behavior often co-occurs with hyperactivity, models were run that removed the effect of hyperactivity from the aggressive behavior variable and aggressive behavior still significantly predicted ASPD, but with lower AUCs of ~0.65 (the opposite scenario of hyperactivity with

aggressive behavior removed was also still able to predict ASPD, but fewer raters were significant). Additionally, in models with multiple raters/externalizing behaviors, hyperactivity was not significant, while aggressive behavior was, and AUCs were not significantly improved over models with only aggressive behavior included. This could be supportive of some studies that have shown only CD (not ADHD) to predict later ASPD [130, 142, 143], reflected in the hyperactivity not being significant in the model with both hyperactivity and aggressive behavior included. Perhaps the relationship of the predictive factors of aggression and hyperactivity with ASPD could be clarified in investigating whether the factors help differentiate those individuals who will have future ASPD and become engaged in criminal activity or not. Our study did not have criminal record information, but it would perhaps be informative since Pingault et al. [212], for example, showed that aggression but not hyperactivity predicted later criminality.

6.2.4 Identifying 3-hydroxybutyrate as a potential biomarker of aggressive behavior (Aim 3)

Using a hypothesis-free NMR approach, Study IV is among only a few studies thus far to examine a panel of (low molecular weight) metabolites in terms of their association with human behavior, and the first to look at aggressive behavior and metabolites in blood. The correlation between aggressive behavior and the metabolites was generally weak (r-values <0.15), and ultimately, the full biomarker panel investigation only indicated a significant (negative) association with 3-hydroxybutyrate and aggressive behavior. This is the first time, to our knowledge, that this metabolite has been found to be related to aggressive behavior, although biomarkers of energy metabolism approached significance in a urine biomarker analysis of aggressive behavior in NTR [168]. When further examined, the negative association trend held among most of aggressive behavior ratings (from eight raters), with fully adjusted models indicating teacher age 12 and self age 14 ratings, in models separately and together, remaining significant. As seen in Study III, the teacher and self ratings appear the most consistently associated with the dependent variable, and each appears to contribute unique information about aggressive behavior, because both remain significant in multi-rater models. This suggests that these two raters hold valuable observational abilities regarding adolescent behavior, and possibly should be seen as raters to prioritize collecting information from in future studies. Of note, the variance in 3-hydroxybutyrate explained by the fully adjusted FT12 models was 1.8–4.2%.

The association of 3-hydroxybutyrate with aggressive behavior in FT12 was also supported in the independent NTR cohort sample. Although the NTR sample utilized aggressive behavior measures in adulthood (not adolescence, like FT12), the general trend still remained, including in Bayesian modeling.

The NTR analysis also indicated a sex interaction (sex x aggression score) in the fully adjusted frequentist model, which was not found in the FT12 frequentist models, but was seen in the Bayesian modeling. The NTR cohort suggested that the negative association was most strongly seen in males, however, in FT12 the trend was stronger in females. Thus, sex differences in this association need to be investigated in more detail in the future. One potential place to look deeper is in subtypes of aggression. Unpublished analyses showed that indirect aggression, in particular, was associated with the biomarker (more consisently than direct aggression), and in Table 8 and 9, males and females had similar indirect aggression levels (age 12 and 14), while boys clearly had higher levels of direct aggression.

Regarding biological plausibility of the association and possible mechanisms, this would remain rather speculative at this point, however, there are some clues. First, 3-hydroxybutyrate is an important part of the alternative energy pathway (when blood glucose is low) and is involved in cell signaling and regulatory functions such as reducing oxidative stress and inflammation (both are associated with aggression; see section 2.1.4 and [213]) and epigenetic modification [214, 215]. Importantly, studies have shown that 3-hydroxybutyrate can cross the blood–brain barrier and act in the brain [216, 217], where behavior is regulated. Additionally, animal and human studies have shown an association of 3-hydroxybutyrate with depression (which can co-occur with aggression), though the direction of the association has been unclear (negative in mice, positive in humans) [214, 218]. Psychiatry has also been investigating the ketogenic diet (which boosts ketone body levels, like 3-hydroxybutyrate), having used it in ADHD and autism, but not yet for aggressive behavior problems [219]. If indeed low 3-hydroxybutyrate is a biological explanation for aggressive behavior, then it would be possible to consider a ketogenic diet (to increase the metabolite's levels) to see if the behavior diminishes. It would also be valuable to investigate 3-hydroxybutyrate alongside insulin and insulin sensitivity/resistance regarding aggressive behavior since 3-hydroxybutyrate is a part of the energy pathway activated when blood glucose is low. However, in our biomarker panel insulin levels were not available, and presumably very few diabetics were in the sample because the FT12 sample was so young.

6.3 Findings in context of the ACTION consortium

This thesis work was performed within the context of collaborative ACTION consortium work, which aimed to provide findings that could improve aggression treatment personalization by enhancing our understanding of the biology, epidemiology, (epi)genetics, and gene x environment interactions in aggression. Studies I–II involved multiple cohorts in the consortium and were

meant to provide valuable phenotype clarity regarding childhood aggressive behavior that could be used to clarify findings from the (epi)genetic (van Dongen et al. bioRxiv https://doi.org/10.1101/2020.07.22.215939; Ip et al. bioRxiv https://doi.org/10.1101/854927; Jami et al. medRxiv https://doi.org/10.1101/2020.09.11.20175026) and biomarker investigations [Study IV; 168]. The ACTION-related work in this thesis involved only individual (between-family) level analyses, but other work in ACTION has also utilized the classical twin modeling approach to estimate the relative importance of genetic and environmental influences on aggression [73, 160]. Additionally, within-family associations have been utilized to investigate if the associations of aggressive behavior with academic performance are confounded by genetic or environmental effects [82]. Many of these cohorts will also likely continue to collaborate in the future, and so the phenotype information will continue to support future research.

6.4 Aggression measurement and reliability

The work in this thesis involved multiple cohorts and instruments, so it is important to consider how some of the instruments relate to one another and how aggressive behavior is captured and investigated using those instruments. The ASEBA instruments are well-known, broadly used and deeply investigated by the creator [5, 102, 148, 151, 220], including investigation of the ASEBA system compared to SDQ [149]. The SDQ is a good short-form for collecting similar information as the ASEBA instruments. In the ACTION consortium, a harmonization effort for some of the cohorts (NTR, TEDS, TCHAD, CATSS, FT12) included three of the study instruments directly (CBCL, SDQ, ATAC) and then the MPNI was incorporated after direct harmonization [160]. Additionally, Hendriks et al. [159] utilized the NTR cohort's CBCL, SDQ, ATAC harmonization data from Luningham et al. [160] and showed that while the instruments have slightly different items, the underlying genetics of the childhood aggression captured by the instruments was consistent. These investigations still leave the MPNI somewhat on the outskirts regarding comparability. However, Study III showed that MPNI could predict a psychiatric disorder, just like the other instruments [148, 221]. Furthermore, in Studies I–II the MPNI produced similar results to the other instruments, and in Study IV the MPNI results were supportively reproduced in the NTR cohort using the ASR. To confirm differences or similarity among the instruments, however, would require systematic testing (e.g., one cohort with multiple instruments predicting the same longitudinal outcome).

Regarding raters of aggressive behavior and reliability of the measurement, it is important to consider how reliable each rater is regarding externalizing behaviors, as well as internalizing behaviors (since Studies I and II look at co-

occurrence). Previous studies – including cohorts used in this thesis – have indicated high reliability for teacher ratings of behavior. For example, by looking at intrapair correlations of ratings of monozygotic (MZ) twins, it is possible to obtain the lower boundary of reliability (note: the reliability could be lower than this if the trait exhibited low heritability, but aggressive behavior was shown to be between 50–80% heritable using NTR and TEDS [73]). This is based on the assumption that the pair-wise correlation in MZ's cannot be greater than test–retest correlation in one individual (see e.g., [222] for MZ correlations providing information about reliability). For example, Polderman et al. [223] used the TRF in NTR and showed that intrapair correlations in MZ pairs on teacher-rated aggressive behavior was 0.84, and between 0.71 and 0.74 for internalizing behavior. Although the intra-pair correlations for internalizing problems were lower than externalizing behavior, the difference was small. Moreover, values of over 0.7 indicate acceptable reliability. Pulkkinen et al. [104] showed that MZ correlations of nearly all MPNI teacher ratings were higher than those of parents (teachers ranged 0.57–0.87, parents ranged 0.43–0.83). Additionally, Verhulst et al. [224] indicated that while teacher ratings for internalizing behavior were less reliable than parent or self ratings, they were still valuable and provided unique information in predicting maladjustment.

Lastly, with aggressive behavior being such a broad and heterogenous behavior, we can also consider which aggressive behavior has been captured in this thesis and its context with other aggressive behavior concepts from Figure 1. The MPNI aggressive behavior scale was separated into direct and indirect aggression subtypes for some of the analyses (Study III and unpublished data from Studies I, II, IV) and was able to suggest nuance to some of the findings (e.g., direct aggression, in particular, predicts ASPD well and indirect aggression, in particular, is associated with 3-hydroxybutyrate). These could provide clues for personalized treatments of aggressive behavior in the future and clarify some of the heterogeneity in the behavior. On the other hand, the MPNI showed similar results in Studies I–II to the ASEBA and SDQ instruments, which both broadly capture aggressive behavior. The ASEBA instruments do not capture indirect aggression, and the SDQ's 'conduct problems' is a mix of aggressive behaviors and rule-breaking behaviors. However, Hendriks et al. [159] showed that these item-level differences still capture similar underlying genetics, and Studies I–II indicate that aggressive behavior often comes along with many co-occuring behaviors. These co-occuring behaviors can all be rolled up and studied together as the p-factor (which combines externalizing and internalizing problems into one general measure of psychopathology), particularly useful when investigating the genetics of psychopathology, as TEDS has done [124].

6.5 Strengths and Limitations

The studies in this thesis had many strengths. All studies utilized large population-based cohorts with high response rates to investigate the research questions. Main findings were able to be supported with replications or by using multiple raters of behaviors. Study I and II shared a very similar protocol so that results could roughly be compared across publications (the children were the same and sample sizes were large), which is of benefit because direct comparison was outside of the scope of both studies. Study III was able to study the main effect of aggressive behavior while removing the residual effect of hyperactivity. Study IV was able to adjust the main findings by multiple confounders similarly across the two cohorts, and suggestively replicate results from FT12 in NTR using a Bayesian approach.

The studies in this thesis did have some notable limitations as well. Study I and II were cross-sectional in nature despite being drawn from longitudinal cohorts. However, the main scope of the studies were to capture co-occuring behavior between the genders at different ages from cohorts using different instruments. Due to the longitudinal nature, many of the same children were captured across the different ages, so there can be an approximate study of patterns across the ages. Regarding the different instruments being compared, although they are different and could not be directly compared, similar patterns across instruments were seen and I previously noted that Hendriks et al. [159] and Luningham et al. [160] showed that the instruments are similar in the underlying genetics they capture. Although Study I did not formally compare the parent and self ratings they presented, and Study I and II ratings were not formally compared, I did provide a supplemental direct comparison of these raters with FT12 data in this thesis.

For Study III, a main limitation for analysis was that the dataset contained only 67 ASPD cases in a starting sample of 1821 participants. However, the 67 cases included both men and women and were in proportion with typical population prevelance of ASPD. The participants were all outside of prison and institutional systems when participating in data collection, so perhaps these ASPD individuals did not fully represent the complete ASPD population. For example, Fazel et al. [136] showed that, across 12 countries, among prison populations 47% of men have ASPD and 21% of women have ASPD; the prison population was 10 times more likely to have ASPD than the general population. However, in Finland, changes in the criminal justice system regarding youth deliquents produced a dramatic reduction in youth in prison between the 1980s and 1990s; although, in the mid-1990s, 17% of youth in prison in Finland had a personality disorder [225]. An additional limitation for Study III is that callous-unemotional traits were not able to be taken into consideration since that type of information was not collected, although they are related to aggressive behavior [226] and ASPD [129].

For Study IV, the main limitation was not having the aggression measures and blood sample taken at the same time (or longitudinally). However, aggression is a rather stable behavior [4, 73], and our samples were limited to people 30 years old and younger, before aggressive behavior really starts to taper off due to age [1]. It would have been helpful to have had repeated measures of the biomarker in order to understand its stability better, however, Nightingale (the company who NMR-processed the samples) does not have such data available yet. Although, for both men and women, 3-hydroxybutyrate is not strongly affected by age [227] and is highly heritable (~55%)[228]. Lastly, because Studies I–III showed that aggressive behavior does not often occur in isolation, hyperactivity or another co-occuring behaviors could have been looked at. However, since this was a new and exploratory approach, the main associations needed to be investigated and replicated first. Future studies should look into the nuances of the association.

CONCLUSIONS

The findings of this thesis contribute to improving our understanding of the heterogenous phenotype of aggression, as well as indicating many gender similarities in the behavior and differences across raters. Study I and II showed that childhood aggression regularly co-occurs with other externalizing behaviors and low prosocial skills and moderately co-occurs with internalizing problems. Patterns were similar across study instrument and age. Most of the variation in aggression was explained by co-occuring externalizing behavior, and co-occurrence of internalizing problems and prosocial behavior with aggression often involves multiple co-occuring behaviors. The co-occurrence findings of Study I and II shed light on the findings in Study III where unadjusted aggression (i.e., the effect of co-occuring behavior(s) could have been influencing the association) predicted ASPD, but even when the effects of co-occuring hyperactivity were removed, aggression still predicted ASPD (and vice versa). This independent effect of aggression in ASPD prediction could help refine DSM characterization (or treatment choice) of the disorder, where currently the presence or absence of callous-unemotional (i.e., psychopathic) traits is the noted important factor when diagnosing and treating ASPD. The co-occurrence of aggression with other behaviors highlights the heterogenous nature of the behavior and the challenges to be encountered if studying aggression without taking other behavioral factors into account.

This thesis revealed several patterns of similarity across gender that should be highlighted against the backdrop of male-dominated aggression research. Boys/Men have long been the group that researchers have focused on for criminal and behavioral problems, and although they exhibit higher levels of these behaviors, Study I and II results have shown that boys and girls both experience co-occurrence to similar degrees. This finding requires researchers and those who refer children for psychological treatment to broaden their focus to all children, and perhaps apply more direct attention to girls' behavioral problems in particular, as the FemNAT-CD consortium has been doing. As with many areas of health research where men have historically been the default gender for initial research investigations, there is much to be gained in aggression research by improving understandings of gender similarities and differences. Indeed, Study III showed that aggressive adolescent behavior in both genders predicted later ASPD, but Study IV suggested possible sex differences in aggression's association with 3-hydroxybutyrate. As cultures and societies continue to evolve, and gender roles and expectations continue to shift and mix, we may see aggression levels change based on both real changes in behavior in the children and reduced biases of observers (i.e., parents and teachers noticing and valuing female

aggressive problems as seriously as male aggressive problems). More research on gender similarities and differences in this heterogenous behavior can continue to help improvement in diagnosis and treatment.

Thesis findings also supported the need for multiple raters of behavior to properly characterize the scope of situational or generalized aggression. Some studies did show some similarities across raters; for example, most ratings of aggressive behavior were suggestively associated with 3-hydroxybutyrate, teacher and self ratings both predicted ASPD, and some teacher–self and parent–teacher direct comparisons of aggressive behavior showed only small effect sizes. On the other hand, Study III and IV showed that when including teacher and self ratings in the same model both can still be significantly associated with ASPD or 3-hydroxybutyrate, indicating both raters contribute some unique insight to the association of interest. Many studies of child behavior utilize parent ratings, however, in Study III and IV in particular, they were the least powerful or consistent for discovering associations. Multiple ratings, especially those beyond the parents, are an important tool in helping to further clarify the aggression phenotype.

Clarifying the aggression phenotype further, this thesis also showed main associations using total aggression were refined when aggression subtypes were also examined. Study III noted that general aggression, and direct aggression in particular, predicted ASPD. Unpublished Study IV analysis presented in the thesis suggests that indirect aggression could be a strong component of the association of aggression with 3-hydroxybutyrate. Furthermore, unpublished analyses in this thesis related to Study I and II suggest that indirect aggression, in particular, differs depending on rater. Teachers see more indirect aggression than parents, but self ratings indicate more indirect aggression than teachers, and this is similar for both genders. While looking at broad characterizations of aggressive behavior (and co-occurrence) can help in investigations across cohorts and instruments, and may capture similar underlying genetic aspects of the behavior, drilling down to details of aggression (e.g., subtypes or gender differences) for investigations of aggressive behavior can help to clarify some of the heterogeneity in the phenotype.

ACKNOWLEDGMENTS

The work for this thesis was performed at the University of Helsinki in the Institute for Molecular Medicine Finland (FIMM) at HiLIFE and the Department of Public Health at the Faculty of Medicine. I wish to thank the present and former heads and directors for providing excellent work facilities and creating welcoming and supportive work environments. My sincere thanks for the funding that allowed this thesis study to be performed: funding awarded to Professor Kaprio from the ACTION and Equal-Life consortiums, as well as an award from the Doctoral Programme in Population Health (DocPop). I must also thank the twins and their families, without whom there would be no data to do this interesting work.

I want to express my gratitude to Professor Essi Viding who agreed to act as the opponent for the public examination of this dissertation and for being flexible and gracious during these ever-changing pandemic times. Additionally, I wish to thank Professors Pirkko Riipinen and Tommi Tolmunen for serving as pre-examiners, providing insightful and detail-oriented comments that certainly improved the thesis and were much appreciated by this fellow detail nerd. And thank you to Professor Anna Keski-Rahkonen for agreeing to be my custos, I feel so lucky to be guided on this final rite-of-passage by you. It has been inspiring to learn from you as a teacher and mentor over the years, and so although you were not an official advisor, I am glad you were able to leave your mark on my journey.

To my thesis committee members, Professor Meike Bartels and Docent Kaisa Riala, I have greatly appreciated your suggestions, support and guidance over the years, thank you for being active and available in your roles. I also wish to thank my many co-authors for helping the publications reach their final destinations. In particular, I wish to thank those patient co-authors who I was able to frequently interact with and learn so much from: Lea Pulkkinen, Dorret Boomsma, Dick Rose, Leonie Bogl, Meike Bartels, Fiona Hagenbeek, Anna But, Anu Raevuori, Koen Bolhuis, and Kaili Rimfeld.

And now I would like to turn to my research family tree, the one that grew up around me without my even noticing until you had rooted yourselves in my life and left me forever changed. When I first had a meeting to join the group (at the Ruskeasuo campus), the meeting table was presided over by a giant poster of Frida Kahlo, and I somehow knew that this was the place for me, and I was not wrong. Thank you to Professor Jaakko Kaprio to being the figurehead of this family. Your deep knowledge of the scientific literature (and history and more) means that our hungry minds always have something to chew on. You seem to have neverending ideas for how to get the most out of

the twin data and have been very generous with encouraging me to travel for meetings and conferences, to collaborate with other research groups, to allow me to occasionally turn my attentions to side projects and ideas, and to give good advice. I look forward to continuing working together. To my advisors, Docents Tellervo Korhonen and Eero Vuoksimaa, I need to thank you for your expert guidance on research aspects, for always being available and patient, and for always holding my wellbeing as an important part of a successful PhD journey. Telle, as a favorite auntie figure and surrogate mummi to my kids (and thank you too Pentti), I have really appreciated your care and attention (including attention to detail and formalities) throughout the thesis work. And Eero, as the big-brother-I-never-had figure, I appreciate you being equal parts willing to share code and give advice on how to prioritize as well as for making me pause and check in on how I'm doing.

To my vibrant collection of helpful, kind, wise, funny colleagues, you fill out the branches of the family tree in a way that makes me feel gathered in the warmth as well as protected. I wish to thank my senior colleagues for their advice, generosity and attention over the years, everywhere I looked there were willing mentors; thank you Antti, Anna, Sari, Leonie, Miina, Maarit, Venla, Karri, Anu L, Emma, Sara L. To my data saviors who were always speedy to deliver data, answer questions, and provide much appreciated support on data matters I thank you so much: Kauko, Teemu, Anja H, Mia U. And to my amazing peers and fellow graduate school survivors, your ears for listening, hearts for sharing, words for advising, and hands for holding warm Nepalese tea at our weekly pilgrimage to our treasured lunchtime locales will always be among my favorite memories of this time of my life. I hope we have many reunions over good food in the future: Mahes, Sailalitha, Anu, Paula, Bram, Jenni, Guiomar, Birgitta, Suvi, Sara, Jade, Eeva-Liisa, Richa, Pauliina, Aino, Bodine, Yasmina.

Beyond my research family, I was so lucky to be able to pick up a charming bunch of unexpected people along my way that helped inspire me, challenge me, entertain me and feed my soul. Thank you to all the students in the courses I taught for teaching me about your work, asking insightful questions, and for coming out of your shells to laugh with me. Thank you to my teacher colleagues who shared stories, ideas, material and laughs. And thank you to all my clients whose academic texts I edited who asked questions that ended up teaching me more and whose work often ended up informing my own.

A long and arduous project such as a PhD would not be so easy to accomplish without caring people outside your academic life to balance things out. I need to thank my friends for keeping me laughing through the pains and reminding me who I am. Thank you Anne for all the walks, talks, laughs, information ;), and kindness over the years. What would I have done without you? Seriously. Thank you Kelsey for embracing the weird upfront and for this wonderful

Michigander connection that has grown into a cherished friendship. Thank you Vanessa for being a wonderful mentor and teacher colleague first, who then came to fill my lunch times with friendship, solidarity, food for thought, and books to feed my soul; and Rachel, when you joined our 'book club', a puzzle piece clicked into place. Thank you for your gentle friendship, explanations and insights. To Lisa, thank you for the willingness to take a chance on me based on a random Facebook message, the world of editing (and then teaching) you opened up for me is such a gift; I wish we could have our lunches more frequently. To Hannah, my first Finnish friend, I am so lucky to have found you and that your love for your country helped me to fall in love with it too. I wish you peace and a way back home soon. To our friends, the Dunns and the Allens, we have so enjoyed our hangouts where the adults can discuss shared experiences in academia and America while the kids can run off and play together. Lastly, to my village back in the US, my cheerleading squad, my maniacal sillies, and my shoulders to cry on: to Amy, whose sword envy drove me to start this PhD in the first place (you next lady!); to my UM SPH friends, our friendships during our masters degree smoothed out some of my rough edges and I love watching your careers and families grow, thanks for the love and support; and to my TK Pod people, whose tolerance for me is beyond human kindness, never change, let's grow old together.

And to my family, the original shapers of who I am, I am grateful for all the time, energy, resources, and love you have shared with me over my years. Family lore suggests that I generally guided my own ship from an early age, but I know that the hard work and attention from my mom and dad were important markers for helping me find my way. Thanks for being my lighthouses. And thank you to my dearly departed in-laws, Mary Ann and Dave, who loved me as their own child ever since I met them as a teenager. Not many people are lucky enough to have two sets of caring parents. Thank you to my brothers and sister-in-law for their love and support. And thank you to my Grandpa, Gran, and Moomoo, who were always happy to see me and share their love. Gran, you always wanted a doctor in the family…I'm not sure if this fully counts, but I'm sure you would have been proud. I miss you.

Lastly, I need to thank Dave for changing the whole course of my life when I glimpsed him on that bus in Vancouver. Your smarts opened up the world to me, brought us to Finland, and the experiences along the way have shown me who I am and who I can be. Thank you for being a soft place for my storms to burn themselves out, for being a mirror to show me all my good bits, and for being my steadfast best, best friend. And to my kids, Dexter and Renn, thank you for breaking me apart and building me anew. I never thought I would or could do a PhD with kids, but you provided (in hindsight) welcome distractions from research and showed me what was truly important. Watching you becoming thoughtful, caring, open-minded people reminds me that you are my best contribution to the world.

REFERENCES

1. Tremblay, R., *The development of aggressive behaviour during childhood: What have we learned in the past century?* Int J Behav Devel, 2000. **24**: p. 124-141.
2. Merikangas, K.R., et al., *Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication--Adolescent Supplement (NCS-A).* J Am Acad Child Adolesc Psychiatry, 2010. **49**(10): p. 980-9.
3. Fergusson, D.M., L.J. Horwood, and E.M. Ridder, *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.
4. Pulkkinen, L., *Longitudinal Study of Personality and Social Development: Insights about aggression after five decades*, in *The Cambridge Handbook of Violent Behavior and Aggression, 2nd Edition*, D.J.F. A. T. Vazsonyi, & M. DeLisi, Editor. 2018, Cambridge University Press: Cambridge, UK. p. 31-51.
5. Achenbach, T.M., S.H. McConaughy, and C.T. Howell, *Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity.* Psychol Bull, 1987. **101**(2): p. 213-32.
6. Connor, D.F., et al., *Gender differences in reactive and proactive aggression.* Child Psychiatry Hum Dev, 2003. **33**(4): p. 279-94.
7. Card, N.A., et al., *Direct and indirect aggression during childhood and adolescence: a meta-analytic review of gender differences, intercorrelations, and relations to maladjustment.* Child Dev, 2008. **79**(5): p. 1185-229.
8. Berkowitz, L., *Aggression: it's causes, consequences, and control.* 1993, New York: McGraw-Hill.
9. Loeber, R. and M. Stouthamer-Loeber, *Development of juvenile aggression and violence. Some common misconceptions and controversies.* Am Psychol, 1998. **53**(2): p. 242-59.
10. Bandura, A., *Aggression: a social learning analysis.* 1973, Englewood Cliffs, NJ: Prentice-Hall.
11. Bettencourt, B.A., et al., *Personality and aggressive behavior under provoking and neutral conditions: a meta-analytic review.* Psychol Bull, 2006. **132**(5): p. 751-777.
12. Buss, A.H. and M. Perry, *The aggression questionnaire.* J Pers Soc Psychol, 1992. **63**(3): p. 452-9.
13. Burt, S.A., *How do we optimally conceptualize the heterogeneity within antisocial behavior? An argument for aggressive versus non-aggressive behavioral dimensions.* Clin Psychol Rev, 2012. **32**(4): p. 263-79.
14. Deater-Deckard, K. and R. Plomin, *An adoption study of the etiology of teacher and parent reports of externalizing behavior problems in middle childhood.* Child Dev, 1999. **70**(1): p. 144-54.
15. Little, T.D., et al., *Disentangling the "whys" from the "whats" of aggressive behaviour.* Int J Behav Devel, 2003. **27**(2): p. 122-133.
16. Turgay, A., *Aggression and disruptive behavior disorders in children and adolescents.* Expert Rev Neurother, 2004. **4**(4): p. 623-32.
17. Merk, M., et al., *The distinction between reactive and proactive aggression: Utility for theory, diagnosis, and treatment?* Eur J Dev Psychol, 2005. **2**(2): p. 197-220.
18. Poulin, F. and M. Boivin, *Reactive and proactive aggression: evidence of a two-factor model.* Psychol Assess, 2000. **12**(2): p. 115-22.
19. Pitkänen, L., *A descriptive model of aggression and nonaggression with applications to children's behaviour.* Jyväskylä studies in education,

psychology and social research,. 1969, Jyväskylä,: Jyväskylän Yliopisto. 208 p.
20. Sanislow, C.A., et al., *Developing constructs for psychopathology research: research domain criteria*. J Abnorm Psychol, 2010. **119**(4): p. 631-9.
21. Dodge, K.A., *The structure and function of reactive and proactive aggression*, in *The development and treatment of childhood aggression*, D.J. Pepler and K.H. Rubin, Editors. 1991, Lawrence Erlbaum Associates, Inc.: Hillsdale, NJ.
22. Blair, R.J., *Neurocognitive models of aggression, the antisocial personality disorders, and psychopathy*. J Neurol Neurosurg Psychiatry, 2001. **71**(6): p. 727-31.
23. Pulkkinen, L., *Proactive and reactive aggression in early adolescence as precursors to anti- and prosocial behavior in young adults*. Aggressive Behavior, 1996. **22**(4): p. 241-257.
24. Archer, J. and S.M. Coyne, *An integrated review of indirect, relational, and social aggression*. Pers Soc Psychol Rev, 2005. **9**(3): p. 212-30.
25. Lagerspetz, K., K. Björkqvist, and T. Peltonen, *Is indirect aggression typical of females? Gender differences in aggressiveness in 11- to 12-year-old children*. Aggress Behav, 1988. **14**: p. 403-414.
26. Kokko, K. and L. Pulkkinen, *Stability of aggressive behavior from childhood to middle age in women and men*. Aggressive Behavior, 2005. **31**(5): p. 485-497.
27. Vaillancourt, T., et al., *Trajectories and predictors of indirect aggression: results from a nationally representative longitudinal study of Canadian children aged 2-10*. Aggress Behav, 2007. **33**(4): p. 314-26.
28. Crick, N.R. and J.K. Grotpeter, *Relational aggression, gender, and social-psychological adjustment*. Child Dev, 1995. **66**(3): p. 710-22.
29. Archer, J., *Sex Differences in Aggression in Real-World Settings: A Meta-Analytic Review*. Review of General Psychology, 2004. **8**(4): p. 291-322.
30. Pulkkinen, L., *Human Development: From Middle Childhood to Middle Adulthood*. 2017, New York: Routledge.
31. Archer, J., *The nature of human aggression*. Int J Law Psychiatry, 2009. **32**(4): p. 202-8.
32. Crick, N.R. and K.A. Dodge, *A review and reformulation of social-information-processing mechanisms in children's social adjustment*. Psychol Bull, 1994. **115**: p. 74-101.
33. Hubbard, J.A., et al., *Reactive and proactive aggression in childhood and adolescence: precursors, outcomes, processes, experiences, and measurement*. J Pers, 2010. **78**(1): p. 95-118.
34. Archer, J., *Does sexual selection explain human sex differences in aggression?* Behav Brain Sci, 2009. **32**(3-4): p. 249-66; discussion 266-311.
35. Wolfer, R. and M. Hewstone, *Intra- Versus Intersex Aggression: Testing Theories of Sex Differences Using Aggression Networks*. Psychol Sci, 2015. **26**(8): p. 1285-94.
36. Nivette, A., et al., *Sex differences in adolescent physical aggression: Evidence from sixty-three low-and middle-income countries*. Aggress Behav, 2018.
37. de Almeida, R.M., et al., *Escalated aggressive behavior: dopamine, serotonin and GABA*. Eur J Pharmacol, 2005. **526**(1-3): p. 51-64.
38. Seo, D., C.J. Patrick, and P.J. Kennealy, *Role of Serotonin and Dopamine System Interactions in the Neurobiology of Impulsive Aggression and its Comorbidity with other Clinical Disorders*. Aggress Violent Behav, 2008. **13**(5): p. 383-395.
39. Hagenbeek, F., et al., *Discovery of biochemical biomarkers for aggression: A role for metabolomics in psychiatry*. Am J Med Genet B Neuropsychiatr Genet, 2016. **171**(5): p. 719-32.

40. Caspi, A., et al., *Role of genotype in the cycle of violence in maltreated children.* Science, 2002. **297**(5582): p. 851-4.
41. Tiihonen, J., et al., *Genetic background of extreme violent behavior.* Mol Psychiatry, 2015. **20**(6): p. 786-92.
42. Bjork, J.M., et al., *Plasma GABA levels correlate with aggressiveness in relatives of patients with unipolar depressive disorder.* Psychiatry Res, 2001. **101**(2): p. 131-6.
43. Archer, J., *The influence of testosterone on human aggression.* Br J Psychol, 1991. **82 (Pt 1)**: p. 1-28.
44. Albert, D.J., M.L. Walsh, and R.H. Jonik, *Aggression in humans: what is its biological foundation?* Neurosci Biobehav Rev, 1993. **17**(4): p. 405-25.
45. Archer, J., *Testosterone and human aggression: an evaluation of the challenge hypothesis.* Neurosci Biobehav Rev, 2006. **30**(3): p. 319-45.
46. Dreher, J.C., et al., *Testosterone causes both prosocial and antisocial status-enhancing behaviors in human males.* Proc Natl Acad Sci U S A, 2016. **113**(41): p. 11633-11638.
47. McCrimmon, R.J., et al., *Anger state during acute insulin-induced hypoglycaemia.* Physiol Behav, 1999. **67**(1): p. 35-9.
48. Goodman, E., et al., *Hostility and adiposity mediate disparities in insulin resistance among adolescents and young adults.* J Pediatr, 2010. **157**(4): p. 572-7, 577 e1.
49. Laederach-Hofmann, K., S. Kupferschmid, and L. Mussgay, *Links between body mass index, total body fat, cholesterol, high-density lipoprotein, and insulin sensitivity in patients with obesity related to depression, anger, and anxiety.* Int J Eat Disord, 2002. **32**(1): p. 58-71.
50. Suarez, E.C., *Sex differences in the relation of depressive symptoms, hostility, and anger expression to indices of glucose metabolism in nondiabetic adults.* Health Psychol, 2006. **25**(4): p. 484-92.
51. Provencal, N., et al., *Childhood chronic physical aggression associates with adult cytokine levels in plasma.* PLoS One, 2013. **8**(7): p. e69481.
52. Zalcman, S.S. and A. Siegel, *The neurobiology of aggression and rage: role of cytokines.* Brain Behav Immun, 2006. **20**(6): p. 507-14.
53. Turnbull, A.V. and C.L. Rivier, *Regulation of the hypothalamic-pituitary-adrenal axis by cytokines: actions and mechanisms of action.* Physiol Rev, 1999. **79**(1): p. 1-71.
54. Barkhudaryan, N. and A.J. Dunn, *Molecular mechanisms of actions of interleukin-6 on the brain, with special reference to serotonin and the hypothalamo-pituitary-adrenocortical axis.* Neurochem Res, 1999. **24**(9): p. 1169-80.
55. Duke, A.A., et al., *Revisiting the serotonin-aggression relation in humans: a meta-analysis.* Psychol Bull, 2013. **139**(5): p. 1148-72.
56. Tremblay, R., et al., *The search for the age of 'onset' of physical aggression: Rousseau and Bandura revisited.* Criminal Behavior and Mental Health, 1999. **9**: p. 24-39.
57. Cote, S.M., et al., *The joint development of physical and indirect aggression: Predictors of continuity and change during childhood.* Dev Psychopathol, 2007. **19**(1): p. 37-55.
58. Pinker, S., *The Better Angels of Our Nature: Why Violence Has Declined.* 2011, New York: Viking Books.
59. Verhulst, F.C. and J. Van der Ende, *The eight-year stability of problem behavior in an epidemiologic sample.* Pediatr Res, 1995. **38**(4): p. 612-7.
60. Broidy, L.M., et al., *Developmental trajectories of childhood disruptive behaviors and adolescent delinquency: a six-site, cross-national study.* Dev Psychol, 2003. **39**(2): p. 222-45.

61. Loeber, R. and D. Hay, *Key issues in the development of aggression and violence from childhood to early adulthood*. Annu Rev Psychol, 1997. **48**: p. 371-410.
62. Vaillancourt, T., et al., *A longitudinal confirmatory factor analysis of indirect and physical aggression: evidence of two factors over time?* Child Dev, 2003. **74**(6): p. 1628-38.
63. Hyde, J., *How large are gender differences in aggression? A developmental meta-analysis*. Dev Psychol, 1984. **20**: p. 722-736.
64. Knight, G.P., R.A. Fabes, and D.A. Higgins, *Concerns about drawing causal inferences from meta-analyses: an example in the study of gender differences in aggression*. Psychol Bull, 1996. **119**(3): p. 410-21.
65. Brame, B., D.S. Nagin, and R.E. Tremblay, *Developmental trajectories of physical aggression from school entry to late adolescence*. J Child Psychol Psychiatry, 2001. **42**(4): p. 503-12.
66. Wildeboer, A., et al., *Early childhood aggression trajectories: Associations with teacher-reported problem behaviour*. International Journal of Behavioral Development, 2015. **39**(3): p. 221-234.
67. Cote, S.M., et al., *The development of physical aggression from toddlerhood to pre-adolescence: a nation wide longitudinal study of Canadian children*. J Abnorm Child Psychol, 2006. **34**(1): p. 71-85.
68. Kokko, K., et al., *Intensity of Aggression in Childhood as a Predictor of Different Forms of Adult Aggression: A Two-Country (Finland and United States) Analysis*. J Res Adolesc, 2009. **19**(1): p. 9-34.
69. Huesmann, L.R., et al., *Stability of aggression over time and generations*. Developmental Psychology, 1984. **20**(6): p. 1120-1134.
70. Conger, R.D., et al., *Angry and aggressive behavior across three generations: a prospective, longitudinal study of parents and children*. J Abnorm Child Psychol, 2003. **31**(2): p. 143-60.
71. Odintsova, V.V., et al., *Genomics of human aggression: current state of genome-wide studies and an automated systematic review tool*. Psychiatr Genet, 2019. **29**(5): p. 170-190.
72. Vierikko, E., et al., *Sex differences in genetic and environmental effects on aggression*. Aggressive Behavior, 2003. **29**(1): p. 55-68.
73. Porsch, R.M., et al., *Longitudinal heritability of childhood aggression*. Am J Med Genet B Neuropsychiatr Genet, 2016. **171**(5): p. 697-707.
74. Moffitt, T.E., *Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy*. Psychol Rev, 1993. **100**(4): p. 674-701.
75. Association, A.P., *Diagnostic and statistical manual of mental disorders*. 5th ed. 2013, Washington, DC: American Psychiatric Association.
76. Raine, A., et al., *Fearlessness, stimulation-seeking, and large body size at age 3 years as early predispositions to childhood aggression at age 11 years*. Arch Gen Psychiatry, 1998. **55**(8): p. 745-51.
77. Sourander, A. and L. Helstela, *Childhood predictors of externalizing and internalizing problems in adolescence. A prospective follow-up study from age 8 to 16*. Eur Child Adolesc Psychiatry, 2005. **14**(8): p. 415-23.
78. Campbell, S.B., D.S. Shaw, and M. Gilliom, *Early externalizing behavior problems: toddlers and preschoolers at risk for later maladjustment*. Dev Psychopathol, 2000. **12**(3): p. 467-88.
79. Farrington, D.P., *Early predictors of adolescent aggression and adult violence*. Violence Vict, 1989. **4**(2): p. 79-100.
80. White, H.R., et al., *Moderators of the dynamic link between alcohol use and aggressive behavior among adolescent males*. J Abnorm Child Psychol, 2013. **41**(2): p. 211-22.
81. Simonoff, E., et al., *Predictors of antisocial personality. Continuities from childhood to adult life*. Br J Psychiatry, 2004. **184**: p. 118-27.

82. Vuoksimaa, E., et al., *Higher aggression is related to poorer academic performance in compulsory education*. J Child Psychol Psychiatry, 2020.
83. Dodge, K.A., et al., *Reactive and proactive aggression in school children and psychiatrically impaired chronically assaultive youth*. J Abnorm Psychol, 1997. **106**(1): p. 37-51.
84. Rohlf, H.L., et al., *Longitudinal Links between Executive Function, Anger, and Aggression in Middle Childhood*. Front Behav Neurosci, 2018. **12**: p. 27.
85. Nagin, D.S. and R.E. Tremblay, *Parental and early childhood predictors of persistent physical aggression in boys from kindergarten to high school*. Arch Gen Psychiatry, 2001. **58**(4): p. 389-94.
86. Tremblay, R., et al., *Physical aggression during early childhood: trajectories and predictors*. Pediatrics, 2004. **114**(1): p. e43-50.
87. Ladd, G.W. and K.B. Burgess, *Charting the relationship trajectories of aggressive, withdrawn, and aggressive/withdrawn children during early grade school*. Child Dev, 1999. **70**(4): p. 910-29.
88. Kokko, K. and L. Pulkkinen, *Aggression in childhood and long-term unemployment in adulthood: a cycle of maladaptation and some protective factors*. Dev Psychol, 2000. **36**(4): p. 463-72.
89. Dubow, E.F., et al., *Middle childhood and adolescent contextual and personal predictors of adult educational and occupational outcomes: a mediational model in two countries*. Dev Psychol, 2006. **42**(5): p. 937-49.
90. Caspi, A., G.H.J. Elder, and D.J. Bem, *Moving against the world: life-course patterns of explosive children*. Dev Psychol, 1987. **23**: p. 308-313.
91. Caspi, A., et al., *Behavioral observations at age 3 years predict adult psychiatric disorders. Longitudinal evidence from a birth cohort*. Arch Gen Psychiatry, 1996. **53**(11): p. 1033-9.
92. Hayatbakhsh, M.R., et al., *Adolescent problem behaviours predicting DSM-IV diagnoses of multiple substance use disorder. Findings of a prospective birth cohort study*. Soc Psychiatry Psychiatr Epidemiol, 2008. **43**(5): p. 356-63.
93. King, S.M., W.G. Iacono, and M. McGue, *Childhood externalizing and internalizing psychopathology in the prediction of early substance use*. Addiction, 2004. **99**(12): p. 1548-59.
94. Timmermans, M., P.A. van Lier, and H.M. Koot, *Which forms of child/adolescent externalizing behaviors account for late adolescent risky sexual behavior and substance use?* J Child Psychol Psychiatry, 2008. **49**(4): p. 386-94.
95. Brook, J.S., et al., *Young adult drug use and delinquency: childhood antecedents and adolescent mediators*. J Am Acad Child Adolesc Psychiatry, 1996. **35**(12): p. 1584-92.
96. Sourander, A., et al., *Who is at greatest risk of adverse long-term outcomes? The Finnish From a Boy to a Man study*. J Am Acad Child Adolesc Psychiatry, 2007. **46**(9): p. 1148-1161.
97. Roff, J.D. and R.D. Wirt, *Childhood aggression and social adjustment as antecedents of delinquency*. J Abnorm Child Psychol, 1984. **12**(1): p. 111-26.
98. Huesmann, L.R., L.D. Eron, and E.F. Dubow, *Childhood predictors of adult criminality: are all risk factors reflected in childhood aggressiveness?* Crim Behav Ment Health, 2002. **12**(3): p. 185-208.
99. Stattin, H. and D. Magnusson, *The role of early aggressive behavior in the frequency, seriousness, and types of later crime*. J Consult Clin Psychol, 1989. **57**(6): p. 710-8.
100. Vitaro, F., et al., *Reactive and proactive aggression differentially predict later conduct problems*. J Child Psychol Psychiatry, 1998. **39**(3): p. 377-85.
101. Pulkkinen, L., *Offensive and Defensive Aggression in Humans - a Longitudinal Perspective*. Aggressive Behavior, 1987. **13**(4): p. 197-212.

102. Achenbach, T.M., *The classification of children's psychiatric symptoms: a factor-analytic study*. Psychol Monogr, 1966. **80**(7): p. 1-37.
103. Achenbach, T.M., et al., *Internalizing/Externalizing Problems: Review and Recommendations for Clinical and Research Applications*. J Am Acad Child Adolesc Psychiatry, 2016. **55**(8): p. 647-56.
104. Pulkkinen, L., J. Kaprio, and R.J. Rose, *Peers, teachers and parents as assessors of the behavioural and emotional problems of twins and their adjustment: the Multidimensional Peer Nomination Inventory*. Twin Res, 1999. **2**(4): p. 274-85.
105. Waschbusch, D.A., *A meta-analytic examination of comorbid hyperactive-impulsive-attention problems and conduct problems*. Psychol Bull, 2002. **128**(1): p. 118-50.
106. Niv, S., et al., *Aggression and Rule-breaking: Heritability and stability of antisocial behavior problems in childhood and adolescence*. J Crim Justice, 2013. **41**(5).
107. Achenbach, T.M., *Manual for the child behavior checklist/2-3 and 1992 profile*. 1992, Burlington, VT: Department of Psychiatry, University of Vermont. xi, 210 p.
108. Rescorla, L., et al., *International epidemiology of child and adolescent psychopathology ii: integration and applications of dimensional findings from 44 societies*. J Am Acad Child Adolesc Psychiatry, 2012. **51**(12): p. 1273-1283 e8.
109. Bartels, M., et al., *Co-occurrence of aggressive behavior and rule-breaking behavior at age 12: multi-rater analyses*. Behav Genet, 2003. **33**(5): p. 607-21.
110. Verhulst, F.C. and J. van der Ende, *"Comorbidity" in an epidemiological sample: a longitudinal perspective*. J Child Psychol Psychiatry, 1993. **34**(5): p. 767-83.
111. Fontaine, N., et al., *Girls' hyperactivity and physical aggression during childhood and adjustment problems in early adulthood: a 15-year longitudinal study*. Arch Gen Psychiatry, 2008. **65**(3): p. 320-8.
112. Vitaro, F., M. Brendgen, and E.D. Barker, *Subtypes of aggressive behaviors: a developmental perspective*. Int J Behav Devel, 2006. **30**(1): p. 1-19.
113. Gresham, F.M., et al., *Comorbidity of hyperactivity-impulsivity-inattention and conduct problems: risk factors in social, affective, and academic domains*. J Abnorm Child Psychol, 1998. **26**(5): p. 393-406.
114. Jester, J.M., et al., *Trajectories of childhood aggression and inattention/hyperactivity: differential effects on substance abuse in adolescence*. J Am Acad Child Adolesc Psychiatry, 2008. **47**(10): p. 1158-65.
115. McGee, R., S. Williams, and P.A. Silva, *Behavioral and developmental characteristics of aggressive, hyperactive and aggressive-hyperactive boys*. J Am Acad Child Psychiatry, 1984. **23**(3): p. 270-9.
116. Garnefski, N. and R.F. Diekstra, *"Comorbidity" of behavioral, emotional, and cognitive problems in adolescence*. J Youth Adolesc, 1997. **26**(3): p. 321-338.
117. Willner, C.J., L.M. Gatzke-Kopp, and B.C. Bray, *The dynamics of internalizing and externalizing comorbidity across the early school years*. Dev Psychopathol, 2016. **28**(4pt1): p. 1033-1052.
118. Eron, L. and L. Huesmann, *The relation of prosocial behavior to the development of aggression and psychopathology*. Aggress Behav, 1984. **10**: p. 201-211.
119. Veenstra, R., et al., *Prosocial and antisocial behavior in preadolescence: Teachers' and parents' perceptions of the behavior of girls and boys*. Int J Behav Devel, 2008. **32**(3): p. 243-251.
120. Hämäläinen, M. and L. Pulkkinen, *Problem behavior as a precursor of male criminality*. Dev Psychopathol, 1996. **8**: p. 443-455.

121. Pulkkinen, L. and R. Tremblay, *Patterns of boys' social adjustment in two cultures and at two different ages: A longitudinal perspective.* Int J Behav Devel, 1992. **15**: p. 527-553.
122. Romano, E., et al., *Multilevel correlates of childhood physical aggression and prosocial behavior.* J Abnorm Child Psychol, 2005. **33**(5): p. 565-78.
123. Caspi, A., et al., *The p Factor: One General Psychopathology Factor in the Structure of Psychiatric Disorders?* Clin Psychol Sci, 2014. **2**(2): p. 119-137.
124. Selzam, S., et al., *A polygenic p factor for major psychiatric disorders.* Transl Psychiatry, 2018. **8**(1): p. 205.
125. Dick, D.M., et al., *Understanding the covariation among childhood externalizing symptoms: genetic and environmental influences on conduct disorder, attention deficit hyperactivity disorder, and oppositional defiant disorder symptoms.* J Abnorm Child Psychol, 2005. **33**(2): p. 219-29.
126. Loeber, R., et al., *Oppositional defiant and conduct disorder: a review of the past 10 years, part I.* J Am Acad Child Adolesc Psychiatry, 2000. **39**(12): p. 1468-84.
127. Burke, J.D., R. Loeber, and B. Birmaher, *Oppositional defiant disorder and conduct disorder: A review of the past 10 years: II.* J Am Acad Child Adolesc Psychiatry, 2002. **41**(11): p. 1275-1293.
128. Maughan, B., et al., *Conduct disorder and oppositional defiant disorder in a national sample: developmental epidemiology.* J Child Psychol Psychiatry, 2004. **45**(3): p. 609-21.
129. Blair, R.J., E. Leibenluft, and D.S. Pine, *Conduct disorder and callous-unemotional traits in youth.* N Engl J Med, 2014. **371**(23): p. 2207-16.
130. Loeber, R., J.D. Burke, and B.B. Lahey, *What are adolescent antecedents to antisocial personality disorder?* Crim Behav Ment Health, 2002. **12**(1): p. 24-36.
131. August, G.J., et al., *Prevalence of ADHD and comorbid disorders among elementary school children screened for disruptive behavior.* J Abnorm Child Psychol, 1996. **24**(5): p. 571-95.
132. Levy, F., et al., *Gender differences in ADHD subtype comorbidity.* J Am Acad Child Adolesc Psychiatry, 2005. **44**(4): p. 368-76.
133. Hofvander, B., et al., *Continuity of aggressive antisocial behavior from childhood to adulthood: The question of phenotype definition.* Int J Law Psychiatry, 2009. **32**(4): p. 224-34.
134. Nock, M.K., et al., *Prevalence, subtypes, and correlates of DSM-IV conduct disorder in the National Comorbidity Survey Replication.* Psychol Med, 2006. **36**(5): p. 699-710.
135. Goldstein, R.B., et al., *The Epidemiology of Antisocial Behavioral Syndromes in Adulthood: Results From the National Epidemiologic Survey on Alcohol and Related Conditions-III.* J Clin Psychiatry, 2017. **78**(1): p. 90-98.
136. Fazel, S. and J. Danesh, *Serious mental disorder in 23000 prisoners: a systematic review of 62 surveys.* Lancet, 2002. **359**(9306): p. 545-50.
137. Cale, E.M. and S.O. Lilienfeld, *Sex differences in psychopathy and antisocial personality disorder. A review and integration.* Clin Psychol Rev, 2002. **22**(8): p. 1179-207.
138. Balsis, S., et al., *An item response theory analysis of DSM-IV personality disorder criteria across younger and older age groups.* Psychol Aging, 2007. **22**(1): p. 171-85.
139. Bucholz, K.K., et al., *A latent class analysis of antisocial personality disorder symptom data from a multi-centre family study of alcoholism.* Addiction, 2000. **95**(4): p. 553-67.
140. Marcus, D.K., et al., *Is antisocial personality disorder continuous or categorical? A taxometric analysis.* Psychol Med, 2006. **36**(11): p. 1571-81.
141. Eronen, M., P. Hakola, and J. Tiihonen, *Mental disorders and homicidal behavior in Finland.* Arch Gen Psychiatry, 1996. **53**(6): p. 497-501.

142. Lahey, B.B., et al., *Predicting future antisocial personality disorder in males from a clinical assessment in childhood*. J Consult Clin Psychol, 2005. **73**(3): p. 389-99.
143. Sourander, A., et al., *Childhood predictors of psychiatric disorders among boys: a prospective community-based follow-up study from age 8 years to early adulthood*. J Am Acad Child Adolesc Psychiatry, 2005. **44**(8): p. 756-67.
144. Lobbestael, J., M. Cima, and A. Arntz, *The relationship between adult reactive and proactive aggression, hostile interpretation bias, and antisocial personality disorder*. J Pers Disord, 2013. **27**(1): p. 53-66.
145. Suris, A., et al., *Measures of aggressive behavior: Overview of clinical and research instruments*. Aggression and Violent Behavior, 2004. **9**(2): p. 165-227.
146. Achenbach, T.M. and L. Rescorla, *Manual for the ASEBA preschool forms & profiles : an integrated system of multi-informant assessment*. 2000, Burlington, VT: ASEBA. xi, 178 p.
147. Achenbach, T.M. and L. Rescorla, *Manual for the ASEBA school-age forms & profiles : an integrated system of multi-informant assessment*. 2001, Burlington, VT: ASEBA. xii, 238 p.
148. Achenbach, T.M., M.Y. Ivanova, and L.A. Rescorla, *Empirically based assessment and taxonomy of psychopathology for ages 1(1/2)-90+ years: Developmental, multi-informant, and multicultural findings*. Compr Psychiatry, 2017. **79**: p. 4-18.
149. Achenbach, T.M., et al., *Multicultural assessment of child and adolescent psychopathology with ASEBA and SDQ instruments: research findings, applications, and future directions*. J Child Psychol Psychiatry, 2008. **49**(3): p. 251-75.
150. Achenbach, T.M., *Commentary: Definitely more than measurement error: but how should we understand and deal with informant discrepancies?* J Clin Child Adolesc Psychol, 2011. **40**(1): p. 80-6.
151. Achenbach, T.M. and C.T. Howell, *Are American children's problems getting worse? A 13-year comparison*. J Am Acad Child Adolesc Psychiatry, 1993. **32**(6): p. 1145-54.
152. Marland, C., et al., *The Autism-Tics, ADHD and other Comorbidities inventory (A-TAC): previous and predictive validity*. BMC Psychiatry, 2017. **17**(1): p. 403.
153. Spivack, G. and M. Levine, *The Devereux Child Behavior Rating Scales: A Study of Symptom Behaviors in Latency Age Atypical Children*. Am J Ment Defic, 1964. **68**: p. 700-17.
154. Pulkkinen, L., et al., *Child behaviour and adult personality: comparisons between criminality groups in Finland and Sweden*. Criminal Behaviour and Mental Health, 2000. **10**(3): p. 155-169.
155. Goodman, R., *The Strengths and Difficulties Questionnaire: a research note*. J Child Psychol Psychiatry, 1997. **38**(5): p. 581-6.
156. Goodman, R., *Psychometric properties of the strengths and difficulties questionnaire*. J Am Acad Child Adolesc Psychiatry, 2001. **40**(11): p. 1337-45.
157. Bucholz, K.K., et al., *A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA*. J Stud Alcohol, 1994. **55**(2): p. 149-58.
158. Hesselbrock, M., et al., *A validity study of the SSAGA--a comparison with the SCAN*. Addiction, 1999. **94**(9): p. 1361-70.
159. Hendriks, A.M., et al., *Content, diagnostic, correlational, and genetic similarities between common measures of childhood aggressive behaviors and related psychiatric traits*. J Child Psychol Psychiatry, 2020.

160. Luningham, J.M., et al., *Harmonizing behavioral outcomes across studies, raters, and countries: application to the genetic analysis of aggression in the ACTION Consortium*. J Child Psychol Psychiatry, 2020. **61**(7): p. 807-817.
161. Achenbach, T.M., et al., *Assessment of adult psychopathology: meta-analyses and implications of cross-informant correlations*. Psychol Bull, 2005. **131**(3): p. 361-82.
162. Ito, W., et al., *BDNF-restricted knockout mice as an animal model for aggression*. Genes Brain Behav, 2011. **10**(3): p. 365-74.
163. Thomas, A.L., S.M. Davis, and H.A. Dierick, *Of Fighting Flies, Mice, and Men: Are Some of the Molecular and Neuronal Mechanisms of Aggression Universal in the Animal Kingdom?* PLoS Genet, 2015. **11**(8): p. e1005416.
164. van Dongen, J., et al., *Epigenome-Wide Association Study of Aggressive Behavior*. Twin Res Hum Genet, 2015. **18**(6): p. 686-98.
165. Kaddurah-Daouk, R. and K.R. Krishnan, *Metabolomics: a global biochemical approach to the study of central nervous system diseases*. Neuropsychopharmacology, 2009. **34**(1): p. 173-86.
166. Gebregiworgis, T. and R. Powers, *Application of NMR metabolomics to search for human disease biomarkers*. Comb Chem High Throughput Screen, 2012. **15**(8): p. 595-610.
167. Emwas, A.H., et al., *NMR Spectroscopy for Metabolomics Research*. Metabolites, 2019. **9**(7).
168. Hagenbeek, F., et al., *Urinary Amine and Organic Acid Metabolites Evaluated as Markers for Childhood Aggression: The ACTION Biomarker Study*. Frontiers in Psychiatry, 2020. **11**: p. 165.
169. Peets, K. and E. Kikas, *Aggressive strategies and victimization during adolescence: Grade and gender differences, and cross-informant agreement*. Aggressive Behavior, 2006. **32**(1): p. 68-79.
170. Loeber, R., et al., *Optimal informants on childhood disruptive behaviors*. Dev Psychopathol, 1989: p. 317-337.
171. De Los Reyes, A. and A.E. Kazdin, *Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study*. Psychol Bull, 2005. **131**(4): p. 483-509.
172. Rettew, D.C., et al., *When parent and teacher ratings don't agree: the Tracking Adolescents' Individual Lives Survey (TRAILS)*. J Child Adolesc Psychopharmacol, 2011. **21**(5): p. 389-97.
173. van der Ende, J. and F.C. Verhulst, *Informant, gender and age differences in ratings of adolescent problem behaviour*. Eur Child Adolesc Psychiatry, 2005. **14**(3): p. 117-26.
174. De Los Reyes, A., et al., *Principles underlying the use of multiple informants' reports*. Ann Rev Clin Psych, 2013. **9**: p. 123-149.
175. Epkins, C.C., *A preliminary comparison of teacher ratings and child self-report of depression, anxiety, and aggression in inpatient and elementary school samples*. J Abnorm Child Psychol, 1993. **21**(6): p. 649-61.
176. Ledingham, J.E., et al., *Agreement among teacher, peer, and self-ratings of children's aggression, withdrawal, and likability*. J Abnorm Child Psychol, 1982. **10**(3): p. 363-72.
177. Stanger, C. and M. Lewis, *Agreement among Parents, Teachers, and Children on Internalizing and Externalizing Behavior Problems*. Journal of Clinical Child Psychology, 1993. **22**(1): p. 107-115.
178. Hudley, C.A., *Comparing teacher and peer perceptions of aggression: An ecological approach*. J Edu Psych, 1993. **85**(2): p. 377-384.
179. Youngstrom, E., R. Loeber, and M. Stouthamer-Loeber, *Patterns and correlates of agreement between parent, teacher, and male adolescent ratings of externalizing and internalizing problems*. J Consult Clin Psychol, 2000. **68**(6): p. 1038-50.

180. Sawyer, M.G., P. Baghurst, and J. Mathias, *Differences between informants' reports describing emotional and behavioural problems in community and clinic-referred children: a research note.* J Child Psychol Psychiatry, 1992. **33**(2): p. 441-9.
181. Clemans, K.H., et al., *Teacher, parent, and peer reports of early aggression as screening measures for long-term maladaptive outcomes: who provides the most useful information?* J Consult Clin Psychol, 2014. **82**(2): p. 236-47.
182. Pulkkinen, L. and T. Pitkänen, *Continuities in aggressive behavior from childhood to adulthood.* Aggressive Behavior, 1993. **19**(4): p. 249-263.
183. Verhulst, F.C., H.M. Koot, and J. Van der Ende, *Differential predictive value of parents' and teachers' reports of children's problem behaviors: a longitudinal study.* J Abnorm Child Psychol, 1994. **22**(5): p. 531-46.
184. Merrell, K.W., R. Buchanan, and O.K. Tran, *Relational aggression in children and adolescents: A review with implications for school settings.* Psychology in the Schools, 2006. **43**(3): p. 345-360.
185. Cairns, R.B. and B.D. Cairns, *Predicting Aggressive Patterns in Girls and Boys - a Developmental-Study.* Aggressive Behavior, 1984. **10**(3): p. 227-242.
186. Doctoroff, G.L. and D.H. Arnold, *Parent-rated externalizing behavior in preschoolers: the predictive utility of structured interviews, teacher reports, and classroom observations.* J Clin Child Adolesc Psychol, 2004. **33**(4): p. 813-8.
187. Boomsma, D., A. Busjahn, and L. Peltonen, *Classical twin studies and beyond.* Nat Rev Genet, 2002. **3**(11): p. 872-82.
188. de Zeeuw, E.L., et al., *Twin specific risk factors in primary school achievements.* Twin Res Hum Genet, 2012. **15**(1): p. 107-15.
189. Pulkkinen, L., et al., *Peer reports of adaptive behavior in twins and singletons: is twinship a risk or an advantage?* Twin Res, 2003. **6**(2): p. 106-18.
190. Bartels, M., et al., *Childhood aggression and the co-occurrence of behavioural and emotional problems: results across ages 3-16 years from multiple raters in six cohorts in the EU-ACTION project.* Eur Child Adolesc Psychiatry, 2018. **27**(9): p. 1105-1121.
191. Kaprio, J., *The Finnish Twin Cohort Study: an update.* Twin Res Hum Genet, 2013. **16**(1): p. 157-62.
192. Kaprio, J., L. Pulkkinen, and R.J. Rose, *Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families.* Twin Res, 2002. **5**(5): p. 366-71.
193. van Beijsterveldt, C.E., et al., *The Young Netherlands Twin Register (YNTR): longitudinal twin and family studies in over 70,000 children.* Twin Res Hum Genet, 2013. **16**(1): p. 252-67.
194. Haworth, C.M., O.S. Davis, and R. Plomin, *Twins Early Development Study (TEDS): a genetically sensitive investigation of cognitive and behavioral development from childhood to young adulthood.* Twin Res Hum Genet, 2013. **16**(1): p. 117-25.
195. Kooijman, M.N., et al., *The Generation R Study: design and cohort update 2017.* Eur J Epidemiol, 2016. **31**(12): p. 1243-1264.
196. Anckarsater, H., et al., *The Child and Adolescent Twin Study in Sweden (CATSS).* Twin Res Hum Genet, 2011. **14**(6): p. 495-508.
197. Lichtenstein, P., et al., *The Swedish Twin study of CHild and Adolescent Development: the TCHAD-study.* Twin Res Hum Genet, 2007. **10**(1): p. 67-73.
198. Verhulst, F.C. and J. van der Ende, *Agreement between parents' reports and adolescents' self-reports of problem behavior.* J Child Psychol Psychiatry, 1992. **33**(6): p. 1011-23.

199. Hudziak, J.J., et al., *Individual differences in aggression: genetic analyses by age, gender, and informant in 3-, 7-, and 10-year-old Dutch twins.* Behav Genet, 2003. **33**(5): p. 575-89.
200. Fergusson, D.M., J.M. Boden, and L.J. Horwood, *Childhood self-control and adult outcomes: results from a 30-year longitudinal study.* J Am Acad Child Adolesc Psychiatry, 2013. **52**(7): p. 709-717 e1.
201. Kokko, K., et al., *Country, sex, and parent occupational status: moderators of the continuity of aggression from childhood to adulthood.* Aggress Behav, 2014. **40**(6): p. 552-67.
202. Jaddoe, V.W., et al., *The Generation R Study: Design and cohort profile.* Eur J Epidemiol, 2006. **21**(6): p. 475-84.
203. Derks, E.M., J.J. Hudziak, and D.I. Boomsma, *Why more boys than girls with ADHD receive treatment: a study of Dutch twins.* Twin Res Hum Genet, 2007. **10**(5): p. 765-70.
204. McConaughy, S.H. and T.M. Achenbach, *Comorbidity of empirically based syndromes in matched general population and clinical samples.* J Child Psychol Psychiatry, 1994. **35**(6): p. 1141-57.
205. Yang, H.J., W.J. Chen, and W.T. Soong, *Rates and patterns of comorbidity of adolescent behavioral syndromes as reported by parents and teachers in a Taiwanese nonreferred sample.* J Am Acad Child Adolesc Psychiatry, 2001. **40**(9): p. 1045-52.
206. Hämäläinen, M., *Aggressive and non-prosocial behaviour as precursors of criminality. - PsycNET.* 1995.
207. Reef, J., et al., *Predicting adult emotional and behavioral problems from externalizing problem trajectories in a 24-year longitudinal study.* Eur Child Adolesc Psychiatry, 2010. **19**(7): p. 577-85.
208. Greenberg, M.T., et al., *The Effects of a Multiyear Universal Social-Emotional Learning Program: The Role of Student and School Characteristics.* Journal of Consulting and Clinical Psychology, 2010. **78**(2): p. 156-168.
209. GH, B., et al., *Prevention effects moderate the association of 5-HTTLPR and youth risk behavior initiation: gene x environment hypotheses tested via a randomized prevention design.* Child development, 2009. **80**(3).
210. Salmivalli, C., A. Kärnä, and P. E, *Counteracting bullying in Finland: The KiVa program and its effects on different forms of being bullied.* International Journal of Behavioral Development, 2011. **35**(5).
211. Schaeffer, C.M., et al., *Modeling growth in boys' aggressive behavior across elementary school: links to later criminal involvement, conduct disorder, and antisocial personality disorder.* Dev Psychol, 2003. **39**(6): p. 1020-35.
212. Pingault, J.B., et al., *Childhood hyperactivity, physical aggression and criminality: a 19-year prospective population-based study.* PLoS One, 2013. **8**(5): p. e62594.
213. Carroll, J.E., et al., *A urinary marker of oxidative stress covaries positively with hostility among midlife community volunteers.* Psychosom Med, 2010. **72**(3): p. 273-80.
214. Setoyama, D., et al., *Plasma Metabolites Predict Severity of Depression and Suicidal Ideation in Psychiatric Patients-A Multicenter Pilot Analysis.* PLoS One, 2016. **11**(12): p. e0165267.
215. Newman, J.C. and E. Verdin, *beta-hydroxybutyrate: much more than a metabolite.* Diabetes Res Clin Pract, 2014. **106**(2): p. 173-81.
216. Guzman, M. and C. Blazquez, *Ketone body synthesis in the brain: possible neuroprotective effects.* Prostaglandins Leukot Essent Fatty Acids, 2004. **70**(3): p. 287-92.
217. Sleiman, S.F., et al., *Exercise promotes the expression of brain derived neurotrophic factor (BDNF) through the action of the ketone body beta-hydroxybutyrate.* Elife, 2016. **5**.

218. Chen, L., Z. Miao, and X. Xu, *beta-hydroxybutyrate alleviates depressive behaviors in mice possibly by increasing the histone3-lysine9-beta-hydroxybutyrylation*. Biochem Biophys Res Commun, 2017. **490**(2): p. 117-122.
219. Bostock, E.C., K.C. Kirkby, and B.V. Taylor, *The Current Status of the Ketogenic Diet in Psychiatry*. Front Psychiatry, 2017. **8**: p. 43.
220. Achenbach, T.M., L.A. Rescorla, and M.Y. Ivanova, *International epidemiology of child and adolescent psychopathology I: diagnoses, dimensions, and conceptual issues*. J Am Acad Child Adolesc Psychiatry, 2012. **51**(12): p. 1261-72.
221. Goodman, R., et al., *Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample*. Br J Psychiatry, 2000. **177**: p. 534-9.
222. Lykken, D.T., *Presidential address, 1981. Research with twins: the concept of emergenesis*. Psychophysiology, 1982. **19**(4): p. 361-73.
223. Polderman, T.J., et al., *Genetic analyses of teacher ratings of problem behavior in 5-year-old twins*. Twin Res Hum Genet, 2006. **9**(1): p. 122-30.
224. Verhulst, F.C., M.C. Dekker, and J. van der Ende, *Parent, teacher and self-reports as predictors of signs of disturbance in adolescents: whose information carries the most weight?* Acta Psychiatr Scand, 1997. **96**(1): p. 75-81.
225. Sailas, E.S., et al., *Mental disorders in prison populations aged 15-21: national register study of two cohorts in Finland*. BMJ, 2005. **330**(7504): p. 1364-5.
226. Waller, R., et al., *Does early childhood callous-unemotional behavior uniquely predict behavior problems or callous-unemotional behavior in late childhood?* Dev Psychol, 2016. **52**(11): p. 1805-1819.
227. Auro, K., et al., *A metabolic view on menopause and ageing*. Nat Commun, 2014. **5**: p. 4708.
228. Kettunen, J., et al., *Genome-wide association study identifies multiple loci influencing human serum metabolite levels*. Nat Genet, 2012. **44**(3): p. 269-76.

APPENDICES

Appendix 1. Modified Multidimensional Peer Nomination Inventory (MPNI)

Questions are formulated to be directed to parents, teachers, or the child answering about themselves or their sibling/twin (original MPNI questions are formulated for peer nomination in a classroom setting). These questions were asked at age 12, 14, 17 for FT12. The English versions of the answer choices are presented here, as adapted from Pulkkinen et al. (1999) and study questionnaires.

To parents: "We ask you to consider Child A…"

To teachers: "Please evaluate the twin and a control student (a student of the same sex as the twin who is immediately after the twin in the alphabet) against each statement…"

To child for self evaluation: "Estimate how well the following descriptions fit you…"

To child for sibling/co-twin evaluation: "Estimate how well the following descriptions fit your twin…"

Answer choices:

3 the characteristic in question fits the child very well, i.e., it is clearly observable in the child

2 the characteristic in question is definitely displayed in the child, but is not as prominent as above

1 the characteristic in question is sometimes displayed in the child, but is not consistent or strong

0 I have not observed the the characteristic in question displayed at all in the child

Questions (organized by subscale, parentheses display the actual order of the item on the MPNI):

Select the degree to which the child fits the described behavior. The child in question…

Behavioral problems

Aggression

Direct aggression
(21) hurts other people when they are angry, e.g., by hitting, kicking, or throwing things at them
(27) calls people names when they are angry with them
(13) teases other people or attacks them for no reason
(25) bullies smaller and weaker kids

Indirect aggression
(18) spreads rumors about other people's personal matters when he/she is mad at them
(9) excludes people from the group, e.g., by saying 'Let's not play/hang around with him/her'

Hyperactivity/impulsivity
(4) is restless, unable to sit still
(37) is hyperactive
(15) is very talkative
(29) is too impatient to wait for their turn
(33) is disobedient at school/home
(11) often acts without thinking about the consequences
(31) runs about and climbs everywhere despite warnings

Inattention
(34; reverse scored) is conscientious with homework
(36) is forgetful
(17) does not listen instructions
(8) cannot concentrate on anything

Emotional problems

Depression
(24) worries a lot
(3) is sad and depressed
(6) is easily offended/starts crying if someone is mean to him/her
(14) is lonely and without friends
(32) clings to adults or is too dependent

Social Anxiety
(10) is shy with other people
(19) is frightened and nervous about new things or new situations

Victimization
(35) gets teased and taunted a lot

Adjustment

Constructiveness
(7) tries to act reasonably even in difficult situations
(20) is considered reliable/trustworthy by most people
(12) is able to sort things out by talking
(26) helps others in need
(16) often stands up for smaller and weaker kids

Compliance
(5) is patient and calm
(28) never argues with others
(2) is friendly to others
(23) avoids difficult situations by doing something else

Social Activity
(30) is popular among friends/classmates
(1) is a good leader
(22) is often with other kids during recess and after school

Appendix 2. ASPD criteria utilized in Study III based on SSAGA interview questions at age 22

Note: ASPD criteria based on DSM-IV ASPD criteria. If individuals had a history of conduct problems before age 15 (asked during age 22 SSAGA interview) and endorsed 3+ ASPD behaviors they were considered to have ASPD.

ASPD criteria (each counts as one point):
*these are asked if they have occurred after age 15

-Started a fight OR has been in a fight at least 3 times
-Teased others
-Harmed animals in a teasing/taunting manner
-Stole money or goods from family or relatives
-Stole from a shop or strangers OR committed forgery (signed someone else's name, used someone else's credit/debit card)
-Broke into someone's house, car, or building
-Stole money by force or threatening
-Ignited fires on purpose
-Damaged property on purpose
-Caused damage to another person on purpose
-Used a gun
-Forced sexual intercourse
-Did not take care of child or family when it was responsibility to do so
-Hit or beat another person
-Was unable to be faithful to a partner for a whole year