

Higher aggression is related to poorer academic performance in compulsory education

Eero Vuoksimaa,¹ Richard J. Rose,² Lea Pulkkinen,³ Teemu Palviainen,¹ Kaili Rimfeld,⁴ Sebastian Lundström,⁵ Meike Bartels,⁶ Catharina van Beijsterveldt,⁶ Anne Hendriks,⁶ Eveline L. de Zeeuw,⁶ Robert Plomin,⁴ Paul Lichtenstein,⁷ Dorret I. Boomsma,⁶ and Jaakko Kaprio^{1,8}

¹Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland; ²Department of Psychological & Brain Sciences, Indiana University, Bloomington, IN, USA; ³Department of Psychology, University of Jyväskylä, Jyväskylä, Finland; ⁴Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK; ⁵Institute of Neuroscience and Physiology, University of Gothenburg, Gothenburg, Sweden; ⁶Department of Biological Psychology, Vrije Universiteit, Amsterdam, The Netherlands; ⁷Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden; ⁸Department of Public Health, University of Helsinki, Helsinki, Finland

Background: To conduct a comprehensive assessment of the association between aggression and academic performance in compulsory education. **Method:** We studied aggression and academic performance in over 27,000 individuals from four European twin cohorts participating in the ACTION consortium (Aggression in Children: Unraveling gene-environment interplay to inform Treatment and Intervention strategies). Individual level data on aggression at ages 7–16 were assessed by three instruments (Achenbach System of Empirically Based Assessment, Multidimensional Peer Nomination Inventory, Strengths and Difficulties Questionnaire) including parental, teacher and self-reports. Academic performance was measured with teacher-rated grade point averages (ages 12–14) or standardized test scores (ages 12–16). Random effect meta-analytical correlations with academic performance were estimated for parental ratings (in all four cohorts) and self-ratings (in three cohorts). **Results:** All between-family analyses indicated significant negative aggression–academic performance associations with correlations ranging from $-.06$ to $-.33$. Results were similar across different ages, instruments and raters and either with teacher-rated grade point averages or standardized test scores as measures of academic performance. Meta-analytical r 's were $-.20$ and $-.23$ for parental and self-ratings, respectively. In within-family analyses of all twin pairs, the negative aggression–academic performance associations were statistically significant in 14 out of 17 analyses ($r = -.17$ for parental- and $r = -.16$ for self-ratings). Separate analyses in monozygotic ($r = -.07$ for parental and self-ratings), same-sex dizygotic (r 's = $-.16$ and $-.17$ for parental and self-ratings) and opposite-sex dizygotic (r 's = $-.21$ and $-.19$ for parental and self-ratings) twin pairs suggested partial confounding by genetic effects. **Conclusions:** There is a robust negative association between aggression and academic performance in compulsory education. Part of these associations were explained by shared genetic effects, but some evidence of a negative association between aggression and academic performance remained even in within-family analyses of monozygotic twin pairs. **Keywords:** Aggression; cognition; development; educational attainment; school performance.

Introduction

Aggression is behaviour with intention to cause harm to others (Anderson & Bushman, 2002). The most common developmental trajectory of physical aggression is a peak in early childhood around 3–4 years of age followed by a gradual decline afterwards (Tremblay, Vitaro, & Cote, 2018). Despite the developmental changes in the average level of physical aggression, children who are consistently highly aggressive can be identified.

Childhood and adolescent aggression are risk factors for many negative outcomes in childhood, adolescence and in adulthood, including criminality, greater risk of substance use disorders and psychiatric disorders (Fergusson, Horwood, & Ridder, 2005b), particularly if childhood aggression continues into adolescence and aggression is high in

intensity (Pulkkinen, 2017; Pulkkinen, 2018). Aggressive children and adolescents are also at risk for school maladjustment, poor school performance and lower educational attainment, and consequently, for lack of occupational alternatives, poorer adjustment to work life, long-term unemployment and social exclusion (Bynner & Parsons, 2002; Fergusson et al., 2005b; Hinshaw, 1992; Kokko & Pulkkinen, 2000).

The mechanisms underlying the negative relationship between aggression and academic performance are not well understood (Hinshaw, 1992). Genetic and to a lesser degree also common environmental influences shared by family members account for a substantial proportion of individual differences in aggression, academic performance and cognitive ability in childhood. (Haworth et al., 2010; Krapohl et al., 2014; Porsch et al., 2016). Twin studies have indicated a shared genetic and environmental effects between aggression and academic performance/

Conflict of interest statement: No conflicts declared.

cognitive ability (Hicks, Johnson, Iacono, & McGue, 2008; Johnson, McGue, & Iacono, 2009; Koenen, Caspi, Moffitt, Rijdsdijk, & Taylor, 2006; Lewis, Asbury, & Plomin, 2017). Such a genetic correlation would also arise under a causal model where one heritable phenotype influences a second heritable phenotype (De Moor, Boomsma, Stubbe, Willemsen, & de Geus, 2008). In addition to bivariate models where standard phenotypic correlation between aggression and academic performance can be decomposed into genetic and environmental correlations, twin data can be used to conduct a quasi-experimental design.

Compared to correlational studies in unrelated individuals, quasi-experimental design of monozygotic (MZ) and dizygotic (DZ) twin pairs is more suitable to address causality questions. Here, it is possible to test if between-family associations can be confirmed in within-family analyses that control for shared genetic (fully in MZs and partly in DZs) and common environmental (in MZs and DZs) effects (Dick, Johnson, Viken, & Rose, 2000; McGue, Osler, & Christensen, 2010). Within-twin-pair analyses can address if the aggression–academic performance association is confounded by third variables (environmental or genetic) even without any measurement of such variables. For example, low parental education may be linked to both poorer offspring academic performance and higher aggression. In studying unrelated individuals, it is not possible to fully control for a third variable that varies between families. However, siblings, including twins, from the same family share parental education making it possible to fully control for the effects of parental education. It is also possible that some of the genetic effects that make people more aggressive also have negative effects on school performance. Again, these kinds of shared effects cannot be fully controlled in a sample of unrelated individuals, noting here that polygenic scores can be used in unrelated samples but these variables capture only a small proportion of genetic variance. Importantly, no large-scale studies have investigated whether within-twin-pair differences in aggression predict within-twin-pair differences in academic performance.

All MZ pairs are of the same sex, but DZ twins can be either from same- or opposite-sex pairs. Like same-sex DZ twins, also opposite-sex DZ twins share on average half of their segregating genes, so these analyses control also in part for genetic influences. It is well established, that on average, males display more aggression than females (Tremblay et al., 2018) and females tend to outperform males in academic performance (Voyer & Voyer, 2014), but it is not known if aggression is associated with academic performance over and above the effect of sex when studying boys and girls from opposite-sex twin pairs. Moreover, is the association between aggression and academic performance similar in boys and girls?

Finally, the issues related to measurement of aggression and academic performance are important in understanding the mechanisms behind the aggression–academic performance relationship. Aggression in children is often rated by parents or teachers and these ratings are only modestly correlated. To date, there are no studies that have systematically studied the associations between aggression and academic performance in the same individuals with different raters. As aggressiveness, at least to some extent, depend on context and age, investigating information from multiple raters and at different ages will yield a better understanding of the effects of aggression on academic performance. Academic performance can be measured either with teacher-rated grade point averages (GPA) or by using standardized tests. When teachers grade their students, it is possible that students' behaviour problems affect their ratings: those with higher aggression receive poorer grades. In contrast, the effects of behavioural problems on the academic performance are not affected by rater bias when using standardized test scores as measures of academic performance.

Here, we performed a comprehensive investigation of the association between aggression and academic performance by implementing both between- and within-family analyses of >27,000 twins from four European countries. These data sets consist of twins from both MZ and DZ pairs and each included at least two different raters (father, mother, self or teacher) of aggression at ages 7–16 and academic performance measures at ages 12–16. Two of the samples included GPAs and two samples had standardized test scores as measures of academic performance.

We had two aims. First, we conducted between-family (individual level) analyses to investigate if the negative aggression–academic performance relationship is evident across ages, raters and academic performance measures in compulsory education. Secondly, we investigated whether these associations can be replicated in within-family (twin-pair level) comparisons that control for shared environmental and genetic influences. These analyses ask: do more aggressive co-twins have poorer academic performance compared to their less aggressive co-twins?

Methods

Participants

The participants of this study were from four twin cohorts participating in the European Union funded Aggression in Children: Unraveling gene-environment interplay to inform Treatment and InterventiON strategies (ACTION) project (Bartels et al., 2018). Data for aggression and academic performance were available for a total of 27,494 individuals including 12,829 full twin pairs. Each cohort included twins

from MZ, same-sex DZ (SSDZ) and opposite-sex DZ (OSDZ) pairs.

Participating cohorts were as follows: (a) FinnTwin12 from Finland (Kaprio, Pulkkinen, & Rose, 2002); (b) the Childhood and Adolescent Twin Study in Sweden (CATSS) (Anckarsater et al., 2011); (c) the Netherlands Twins Registry (NTR) (van Beijsterveldt et al., 2013); and (d) Twins Early Development Study (TEDS) from the UK (Haworth, Davis, & Plomin, 2013). General descriptions about participant recruitment and study protocols can be found from the references above and from the ACTION consortium website (<http://www.action-euproject.eu/>). Here, we give a short description of each cohort and details about the phenotypes. We used parental (FinnTwin12, CATSS, NTR and TEDS), teacher (FinnTwin12 and NTR) and self-ratings (FinnTwin12, CATSS, NTR and TEDS) of aggression. Academic performance measures were either teacher-rated GPA (FinnTwin12 and CATSS) or standardized test performance (NTR and TEDS).

FinnTwin12. The participants were from a population-based longitudinal FinnTwin12 study that comprises Finnish twins born 1983–1987 (Kaprio et al., 2002). Initially, all twins from these five birth cohorts were identified through Finnish Population Register Centre. Twins and their parents were enrolled in the study when twins were 11–12 year olds with a baseline participation rate above 90%. Study protocols in FinnTwin12 are approved by the Helsinki and Uusimaa Hospital District ethical review board, Helsinki, Finland and Institutional Review Board of the Indiana University, Bloomington, USA. Parents gave permission for assessing their twin children and contacting the teachers of the twins.

CATSS. The CATSS sample consists of twins born from January 1994 and onwards. All twins in Sweden are identified via the Swedish Twin Registry and their parents are invited to participate in a telephone interview in connection with their twins' 9th birthday. Parents and twins are then contacted again when the twins turn 15 and are asked to, separately, fill out a web-questionnaire. The CATSS is described in detail in (Anckarsater et al., 2011). CATSS has ethical permission from the regional ethical review board in Stockholm and at age 15 twins and parents gave separate consents.

NTR. NTR recruits newborn twins and multiples from the population in the Netherlands who were born after 1986 (van Beijsterveldt et al., 2013; Boomsma et al., 2002). Parents receive surveys after birth of twins for every 2–3 years, until the children are 12 years of age. At twins' ages 7, 9/10 and 12 years, the parents are asked for consent to approach the teacher of the twins. After consent was given, teachers were approached with a request to take part in the study. NTR studies were approved by the Central Ethics Committee on Research Involving Human Subjects, VU University Medical Center, Amsterdam.

TEDS. Participants from the population-based study in the UK were drawn from the Twins Early Development Study (TEDS) (Haworth et al., 2013; Oliver & Plomin, 2007; Rimfeld et al., 2019). TEDS is a large longitudinal study that recruited over 16,000 twin pairs born in England and Wales between 1994 and 1996. Although there has been some attrition, more than 10,000 twin pairs remain actively involved in the study. Rich cognitive and behavioural data, including educational achievement, have been collected from the twins, their parents and teachers, over compulsory education and beyond. Ethical approval for TEDS was received from King's College London Ethics Committee. Written informed consent was obtained from participants and/or their parents at each wave of data collection.

Aggression measures

Aggression was measured with the Dutch version of the Child Behavior Checklist (CBCL) and the Teacher Report Form (TRF) in the NTR with mother, father and teacher ratings at ages 7, 10 and 12 years (Achenbach, Ivanova, & Rescorla, 2017; Hudziak et al., 2003). Parental ratings included 18 items each rated in a three-point scale: 0 = behaviour not true, 1 = behaviour somewhat or sometimes true, 2 = behaviour very or often true (items are shown in Table S1). Teacher ratings from the Teacher Report Form (TRF) included 20 items with similar scoring (one item, disobedient at home was excluded and three additional questions asked about behaviour at school). Items were scored on a 3-point scale [range between 0 and 2], and the mean item score was used as an index of aggression.

In FinnTwin12, aggression was measured using Multidimensional Peer Nomination Inventory (MPNI) with parental and teacher ratings at age 12 and self- and teacher ratings at age 14 (Pulkkinen, Kaprio, & Rose, 1999). Both teacher and school at age 14 differed from those at age 12. A total of six questions about aggressive behaviour (four questions about direct aggression and two questions about indirect aggression) were rated on four-point scales: 0 = does not apply, 1 = applies sometimes, but not consistently, 2 = certainly applies, but not in a pronounced way, 3 = applies in a pronounced way (items are shown in Table S1). A mean score of items [range 0 to 3] was used as a measure of aggression.

The Conduct Problem Scale from the Strength and Difficulties Questionnaire (SDQ) was used as a proxy measure of aggression in CATSS (at age 15) and TEDS (at age 16). Both CATSS and TEDS had parental and self-ratings. Each item was rated using three-point scale: 0 = not true, 1 = somewhat true, 2 = certainly true (items are shown in Table S1). A mean score of 5 items ranging from 0 to 2 formed a conduct problems scale that was used as a proxy for aggressive behaviour.

Academic performance measures

Grade point averages. Academic performance was measured with teacher-rated GPA in FinnTwin12 at ages 12 and 14 and in CATSS at age 15. In the Finnish school system, each subject is graded from 4 to 10 and together all subjects yield a GPA ranging from 4 to 10 (Latvala et al., 2014). In Sweden, The National School Register includes grades in all subjects for all students at the end of the 9th grade; these data were merged with the CATSS data. Each subject is graded from 0 to 20 (A = 20 points, B = 17.5 points, C = 15 points, D = 12.5 points, E = 10 points and F = 0 points), and 16 best grades are summed to yield a GPA ranging from 0 to 320.

Standardized educational attainment (EA) test scores. Academic performance was operationalized as a result of a standardized test score. In the NTR, we analysed Dutch CITO-elementary test score at age 12 (de Zeeuw, van Beijsterveldt, Glasner, de Geus, & Boomsma, 2016), and in TEDS, we used General Certificate of Secondary Education (GCSE) standardized examination score at age 16.

CITO test. The CITO test is a standardized test for educational achievement that is administered in the final grade (when children are 11 or 12 years old) of elementary school (Eindtoets Basisonderwijs, 2002). The CITO test consists of multiple choice items in four different educational skills, namely Arithmetic, Language, Study Skills, and Science and Social Studies. The first 3 test scales are combined to a standardized total score (also called CITO score), which is standardized on scale from 500 to 550. Initially, NTR obtained the CITO scores from surveys mailed to teachers of twins. Because results are only available near the very end of the

school year, we later asked the parents to report on the scores. For a small group, the CITO scores were obtained by linking our database of CITO.

GCSE examination. GCSE examination results were obtained from twins themselves or from their parents via questionnaires sent over mail or via telephone. GCSEs are UK-wide standardized examinations taken at age of 16 at the end of compulsory education. Children choose from variety of different subjects from English, math and science, to history, art and geography, and take around 10 GCSEs on average. English, Mathematics and Science are compulsory subjects, so we analysed examination grades from English, Mathematics and Science. Composite measures were created for English (mean of English language and English literature grades), Science (mean of single or double-weighted science, or, when taken separately chemistry, physics and biology grade) and Mathematics. Mean core academic achievement was calculated by taking the mean of English, Mathematics and Science (Krapohl et al., 2014; Shakeshaft et al., 2013). For 7,367 twins the grades were verified in the National Pupil Database (NPD; https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/251184/SFR40_2013_FINALv2.pdf), yielding a correlation of 0.99 for mathematics, 0.98 for English and >0.95 for all the sciences (Rimfeld, Kovas, Dale, & Plomin, 2015).

Data analysis

Between-family association analyses. We used linear regression analyses with academic performance (GPA or test score) as a dependent variable and aggression, sex and age as independent variables. Non-independence of individuals, that is, twins within families, was taken into account in between-family analyses (robust standard errors adjusted for clustered family data in Stata). The age variable was centred, sex coded as 0 = boy, 1 = girl. Constants in these models indicate academic performance in boys with an average age of the sample and no aggression. The regression coefficient in all models indicates change in academic performance units as a function of one unit increase in the aggression score thus allowing direct country-specific interpretation of the effect of aggression on academic performance.

Within-family association analyses. We used fixed effect linear conditional regression analyses to investigate if the associations between aggression and academic performance can be replicated in within-family comparison of twin pairs (xtreg with option fe in Stata). First, analyses included all twin pairs followed by separate analyses including only MZ pairs, DZ pairs from same-sex pairs and DZ pairs from opposite-sex pairs. In line with between-family analyses, centred age and sex (boys as a reference group) were included as covariates in these analyses.

Controlling for shared genetic and environmental effects does not require information about measured genes or environmental factors. Indeed, common environmental effects refer to all environmental influences that make twins within a pair alike, in addition to shared genetic influences. Shared genetic effects are controlled, because MZ twins are genetically identical. Shared genetic effect are also controlled, in part, in DZ twins who on average, like non-twin full siblings, share half of their segregating genes.

Results of within-family analyses clarify whether the associations are confounded by genetic or environmental effects (Dick et al., 2000; McGue et al., 2010). If within-twin-pair differences in aggression are related to within-twin-pair differences in academic performance equivalently in both DZ and MZ pairs and if these within-family correlations are of similar magnitude and equal that found in between-family (individual

level) correlations, there is no third variable confounding. Evidence of confounding is apparent when within-twin-pair aggression–academic performance correlations are smaller in magnitude than individual level correlations. More specifically, partial genetic confounding occurs when the aggression–academic performance correlation within DZ pairs is less than that of between-family correlation, and the correlation within MZ pairs is about half of the correlation within DZ pairs. Complete genetic confounding occurs when the aggression–academic performance correlation within DZ pairs is about half of the individual level correlation and correlation within MZ pairs is (near or) at zero. Complete environmental confounding occurs if both, within DZ and MZ pair aggression–academic performance correlations, are zero.

Meta-analytical correlations. We performed random effects meta-analytical correlations by using metacor function of “meta” package in R (Schwarzer, 2007). Effects were converted to Fisher’s z-transformation weighted by their inverse variances. Sample heterogeneity was assessed with I^2 -statistics, and between-study variance τ^2 were estimated by DerSimonian–Laird method. Meta-analytical approach was used separately for different raters that were included at least in three cohorts. Parental ratings were included in all four cohorts. In the NTR, there were separate aggression ratings from mothers and fathers, but we included only maternal ratings in the meta-analysis and combined them with other cohorts’ parental ratings (other cohorts had a single parental rating, but in a majority of cases, these were based on maternal reports). Self-ratings were included in three cohorts (FinnTwin12, CATSS and TEDS). Further, in meta-analyses, we included only correlations that included aggression rating and academic performance from the same age. In the NTR, maternal ratings were available from different ages but we included only age 12 ratings.

Results

Sample demographics

Table 1 shows mean age at the time of aggression assessments and the number of participants with aggression and academic performance measures for each age for each cohort. The number of twins with available aggression or academic performance data by sex can be found from Tables S2 and S3.

Sex differences in aggression

Boys had higher levels of aggression in FinnTwin12, NTR and TEDS, whereas in CATSS boys had higher aggression based on self-ratings but girls had higher aggression based on parental ratings (Figure 1A, Table S2). The effect sizes of sex differences were modest when measured with CBCL, TRF (in NTR) and MPNI (in FinnTwin12), whereas the effect size of sex differences was negligible or small when measured with SDQ (in CATSS and TEDS) (Figure 1). Compared to parental ratings, teacher ratings yielded consistently greater sex differences (Figure 1).

Sex differences in academic performance

In Finland and Sweden, girls had significantly higher teacher-rated GPAs than boys (Figure 1B, Table S3).

Table 1 Mean age (SD in parentheses) in aggression measurements and number of participants with aggression and academic performance measures in four studies

	7 years	10 years	12 years	14 years	15 years	16 years	
NTR (CITO score at 12 years)	M: 7.40 (0.39), N = 9,812 F: 7.40 (0.39), N = 7,599 T: 7.48 (0.34), N = 4,355	M: 10.03 (0.41), N = 9,071 F: 10.03 (0.41), N = 6,777 T: 10.01 (0.47), N = 5,381	M: 12.26 (0.40), N = 10,388 F: 12.26 (0.40), N = 7,578 T: 12.15 (0.29), N = 5,305	P: 11.79 (0.29), N = 2,884 T: 11.58 (0.32), N = 3,081	S: 14.19 (0.15), N = 1,287 T: 14.23 (0.20), N = 2,828	S: 15.62 (0.48), N = 5,288 P: 15.58 (0.49), N = 4,488	
FinnTwin12 (GPA at 12 years and 14 years)							
CATSS (GPA at 15 years)							
TEDS (GCSE at 16 years)						S: 16.31 (0.68), N = 8,695 P: 16.31 (0.68), N = 8,720	

FinnTwin12, Finnish twin cohort study; Children and Adolescent Twin Study in Sweden; GCSE, examination results for the General Certificate of Secondary Education; NTR, The Netherlands Twin Registry; TEDS, Twins Early Development Study. F, paternal reported; M, maternal reported; P, parental reported; S, self-reported; T, teacher reported.

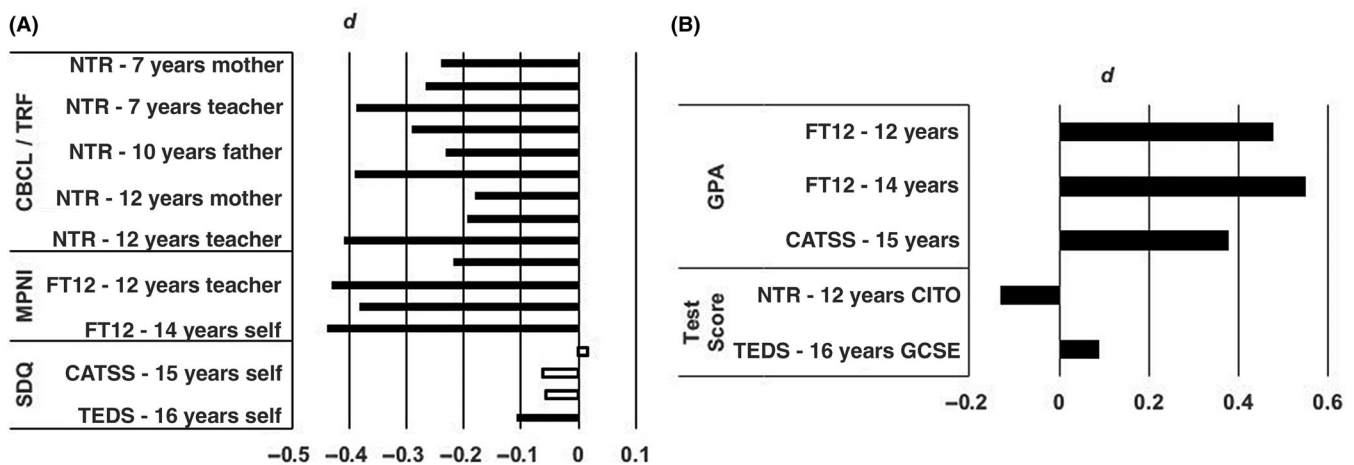


Figure 1 (A) Effect size of sex on aggression as indicated by Cohen's *d*. Negative values indicate higher aggression in boys. Black bars indicate significant sex difference after Bonferroni correction ($p < .003$ [0.05/17]). (B) Effect size of sex on academic performance as indicated by Cohen's *d*. Positive values indicate female advantage. All sex differences were significant at $p < .001$. FT12, FinnTwin12; CATSS, Children and Adolescent Twin Study in Sweden; NTR, Netherlands Twin Register; TEDS, Twins Early Development Study; CBCL, Child Behavior Checklist; TRF, Teacher Report Form; MPNI, Multidimensional Peer Nomination Inventory; SDQ, Strengths and Difficulties Questionnaire. GPA, teacher-rated grade point average. CITO, Dutch CITO-elementary test score; GCSE, General Certificate of Secondary Education examination results

The effect sizes of these differences were modest. Considering standardized EA test scores, girls had higher GCSE scores than boys in TEDS whereas boys had higher CITO scores than girls in the Netherlands (Figure 1B, Table S3). Sex differences in standardized test scores were small in magnitude.

Between-family analyses of aggression and academic performance associations

All between-family analyses indicated significant negative associations between aggression and academic performance across different ages, instruments and

Table 2 Linear regression analyses and the correlations between aggression and academic performance. Unstandardized regression coefficients (with 95% confidence intervals in parentheses) indicate study specific effects of one unit increase in aggression scale on academic performance

Cohort					
Age/Rater	N	Test statistics	b (95% CI's)	t	r
NTR					
Age 7—mother	9,812	F(3, 5383) = 44.06	-3.96 (-4.78; -3.13)	-9.44*	-.10*
Age 7—father	7,599	F(3, 4133) = 32.28	-3.86 (-4.86; -2.86)	-7.55*	-.09*
Age 7—teacher	4,355	F(3, 2498) = 15.20	-3.42 (-5.00; -1.85)	-4.27*	-.06*
Age 10—mother	9,071	F(3, 4985) = 49.13	-4.89 (-5.78; -4.00)	-10.78*	-.13*
Age 10—father	6,777	F(3, 3690) = 35.40	-4.65 (-5.75; -3.55)	-8.29*	-.11*
Age 10—teacher	5,381	F(3, 3090) = 40.76	-5.67 (-6.90; -4.44)	-9.04*	-.12*
Age 12—mother	10,388	F(3, 5689) = 65.66	-5.30 (-6.17; -4.43)	-11.92*	-.13*
Age 12—father	7,578	F(3, 4122) = 45.76	-5.08 (-6.16; -4.00)	-9.20*	-.11*
Age 12—teacher	5,305	F(3, 3094) = 35.32	-6.46 (-7.88; -5.04)	-8.94*	-.13*
FinnTwin12					
Age 12—parent	2,884	F(3, 1509) = 62.14	-0.23 (-0.31; -0.16)	-6.01*	-.16*
Age 12—teacher	3,081	F(3, 1611) = 95.65	-0.25 (-0.30; -0.21)	-10.67*	-.26*
Age 14—self	1,287	F(3, 728) = 37.32	-.29 (-0.43; -0.16)	-4.24*	-.18*
Age 14—teacher	2,828	F(3, 1630) = 139.24	-0.50 (-0.58; -0.43)	-13.42*	-.33*
CATSS					
Age 15—self	5,288	F(3, 2911) = 146.38	-46.43 (-52.34; -40.51)	-15.39*	-.26*
Age 15—parent	4,488	F(3, 2378) = 114.73	-56.35 (-64.40; -48.31)	-13.73*	-.25*
TEDS					
Age 16—self	8,695	F(3, 4464) = 123.09	-0.93 (-1.03; -0.84)	-19.09*	-.22*
Age 16—parent	8,720	F(3, 4434) = 134.21	-1.17 (-1.29; -1.06)	-19.88*	-.26*

All models are adjusted for age, sex and family structure. NTR, Netherlands Twin Registry; CATSS, Childhood and adolescent Twin Study in Sweden; TEDS, Twins Early Development Study; Academic performance is measured with standardized test scores in NTR and TEDS, and grade point average in FinnTwin12 and CATSS; * $p < .001$. All correlations are significant at $p < .05$ using Bonferroni correction for multiple testing.

raters with correlations ranging from -0.06 to -0.33 (Table 2; Figure 2).

Meta-analytical parental-rated aggression–academic performance correlation was -0.20 and self-rated aggression–academic performance correlation was -0.23 (Figure 3A,B). There was significant heterogeneity between parental-rated studies: $Q(3) = 105.98$, $p < .001$, $I^2 = 97\%$, $\tau^2 = 0.0056$ (Figure 3A). There was also significant heterogeneity between self-rated studies: $Q(2) = 9.84$, $p = .007$, $I^2 = 79.7\%$, $\tau^2 = 0.0009$ (Figure 3B).

Sex effects on academic performance remained when controlling for aggression (Tables S4–S9). Aggression by sex interaction effects on academic performance were non-significant in most of the models; only FinnTwin12 age 12 parental ($p = .017$) ratings indicated stronger aggression–academic performance associations in boys than in girls and the NTR age 12 father ratings indicated stronger association in girls ($p = .022$) (correlations by sex are shown in the Table S10).

Within-family analyses of aggression and academic performance associations

In within-family analyses of all twin pairs, the negative association between aggression and academic performance was statistically significant in 14 out of 17 analyses (number of full pairs ranging from 558 to 4,698) (Figure 4, Table S11). Correlations between twin-pair differences in aggression and

twin-pair differences in academic performance ranged from -0.01 to -0.23 (Table S12). Meta-analytical within-twin-pair parental-rated aggression–academic performance correlation was -0.17 ($p < .001$) and within-twin-pair self-rated aggression–academic performance correlations was -0.16 ($p < .001$) (Figure 3C,D). There was significant heterogeneity between parental-rated studies: $Q(3) = 33.28$, $p < .001$, $I^2 = 91\%$, $\tau^2 = 0.0035$ (Figure 3C). There was also significant heterogeneity between self-rated studies: $Q(2) = 9.56$, $p = .008$, $I^2 = 79.1\%$, $\tau^2 = 0.0020$ (Figure 3D).

In MZ pairs, only 3 out of 17 analyses (number of full pairs from 217 to 1,789) indicated a significant negative association between aggression and academic performance. However, meta-analytical MZ within-twin-pair parental-rated aggression–academic performance correlation was $-.07$ ($p < .001$) and MZ within-twin-pair self-rated aggression–academic performance correlation was $-.07$ ($p = .020$) (Figures S1 and S2). There was no significant heterogeneity between parental-rated studies: $Q(3) = 0.10$, $p = .992$, $I^2 = 0\%$, $\tau^2 = 1.00$ (Figure S1). Similarly, there was no significant heterogeneity between self-rated studies: $Q(2) = 3.83$, $p = .147$, $I^2 = 47.8\%$, $\tau^2 = 0.0015$ (Figure S2).

In DZ's from same-sex pairs, 9 out of 17 analyses (number of full pairs from 172 to 1,433) indicated significant negative associations and in DZ's from opposite-sex pairs, 8 out of 17 analyses (number of full pairs from 165 to 1,476) indicated significant

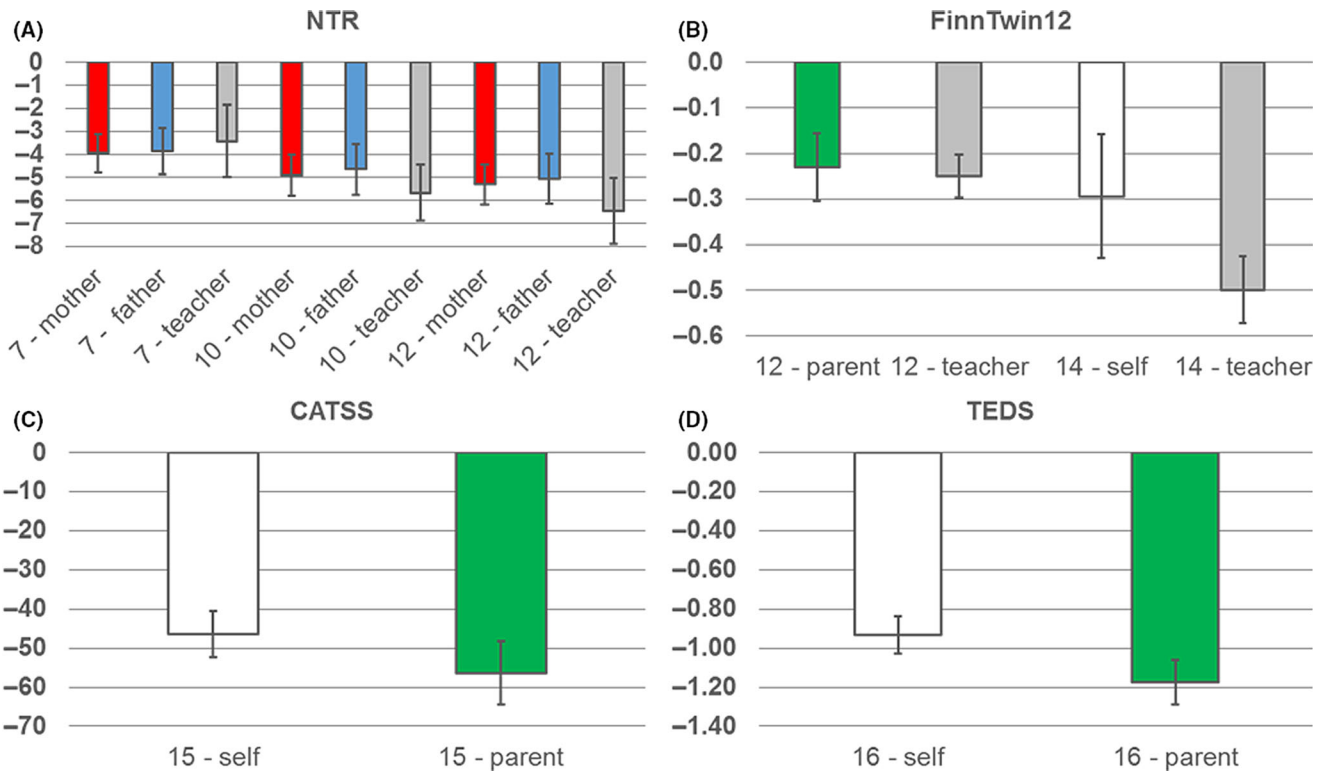


Figure 2 Between-family analyses on the association between aggression and academic performance. Unstandardized regression coefficient (with 95% confidence intervals) indicate the effect of one unit increase in the aggression measure on academic performance in a model that controlled for centred age and sex (girls as reference). NTR, Netherlands Twin Registry; FinnTwin12, Finnish Twin Cohort study; CATSS, Child and Adolescent Twin Study in Sweden; TEDS, Twins Early Development Study. FinnTwin12 and CATSS had teacher-rated grade point average as dependent variables and NTR and TEDS had standardized test scores as dependent variables. See Tables S3–S8 for details [Colour figure can be viewed at wileyonlinelibrary.com]

negative associations between aggression and academic performance (Tables S11 and S12). Generally, the within-family associations became weaker as a function of controlling for greater genetic relatedness (Figure 4, Table S11). Meta-analytical within-twin-pair parental-rated aggression–academic performance correlations were -0.16 and -0.21 for SSDZ’s and

OSDZ (p ’s $< .001$), respectively (Figures S3 and S4). Within-twin-pair self-rated aggression–academic performance correlations were -0.17 and -0.19 for SSDZ’s and OSDZ’s (p ’s $< .001$), respectively (Figures S5 and S6). In SSDZ, there was significant heterogeneity between parental-rated studies: $Q(3) = 13.19$, $p = .004$, $I^2 = 77.3\%$, $\tau^2 = 0.0038$

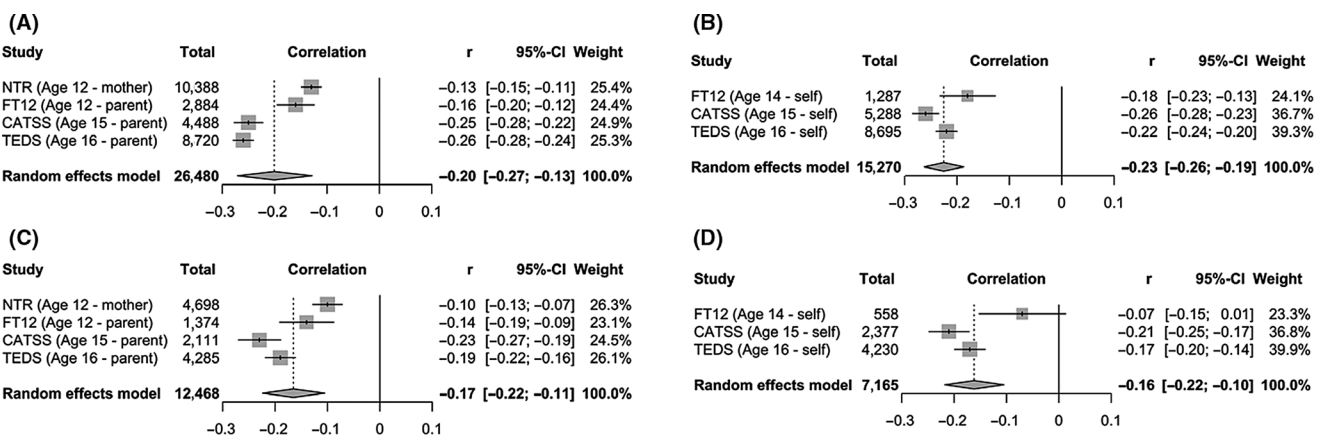


Figure 3 Meta-analysis of between-family (individual level) and within-family (twin-pair level) correlations between aggression and academic performance. Random effect estimates with 95% confidence intervals (CI). (A) Meta-analysis of between-family parental-rated aggression – academic performance correlations. (B) Meta-analysis of between-family self-rated aggression – academic performance correlations. (C) Meta-analysis of within-family parental-rated aggression – academic performance correlations. (D) Meta-analysis of within-family self-rated aggression – academic performance correlations. NTR, Netherlands Twin Registry; FinnTwin12, Finnish Twin Cohort study; CATSS, Child and Adolescent Twin Study in Sweden; TEDS, Twins Early Development Study

(Figure S3), but no significant heterogeneity between self-rated studies: $Q(2) = 5.47, p = .065, I^2 = 63.4\%, \tau^2 = 0.0030$ (Figure S4). In OSDZ, there was significant heterogeneity between parental-rated studies: $Q(3) = 45.09, p < .001, I^2 = 93.3\%, \tau^2 = 0.0156$ (Figure S5), and also significant heterogeneity between self-rated studies: $Q(2) = 8.11, p = .017, I^2 = 75.3\%, \tau^2 = 0.0051$ (Figure S6).

Discussion

A negative relationship of externalizing behaviour with academic performance has been established fairly consistently (Hinshaw, 1992), but the underlying mechanisms of that association are not well understood. Here, we performed a comprehensive assessment on the association between aggression and academic performance in compulsory school aged children/adolescents by using different aggression measures and raters. We also investigated whether the association between aggression and academic performance can be confirmed in a within-family design controlling for genetic and common environmental confounding, analysing data of MZ and DZ twin pairs. Analyses were carried out in twins from four cohorts from four countries participating in the ACTION consortium with aggression measures from three different instruments

rated by parents, self and teachers at ages 7–16. Academic performance at ages 12–16 was measured with grade point averages or standardized test performances.

Sex differences in aggression and academic performance were mostly in expected directions (Tremblay et al., 2018; Voyer & Voyer, 2014); boys had higher aggression than girls consistently across development, countries and raters (with the only exception being parental-rated aggression in CATSS). However, the SDQ scales yielded negligible sex differences. SDQ includes less items compared to CBCL and less answering options compared to MPNI; these factors may have resulted in a smaller sex difference. On the other hand, SDQ was used in CATSS and TEDS at ages 15 and 16, respectively, whereas other cohorts measured aggression in younger ages from 7 to 14 years. Girls outperformed boys in academic performance when using teacher-rated GPA. Standardized test scores yielded sex differences in opposite directions (higher CITO scores in boys versus higher GCSE scores in girls) but these differences were small in magnitude. Standardized tests, like general cognitive ability or g-factor, combine different cognitive domains, and if appropriately balanced with sex neutral, male favouring and female favouring tests; should not show any sex difference.

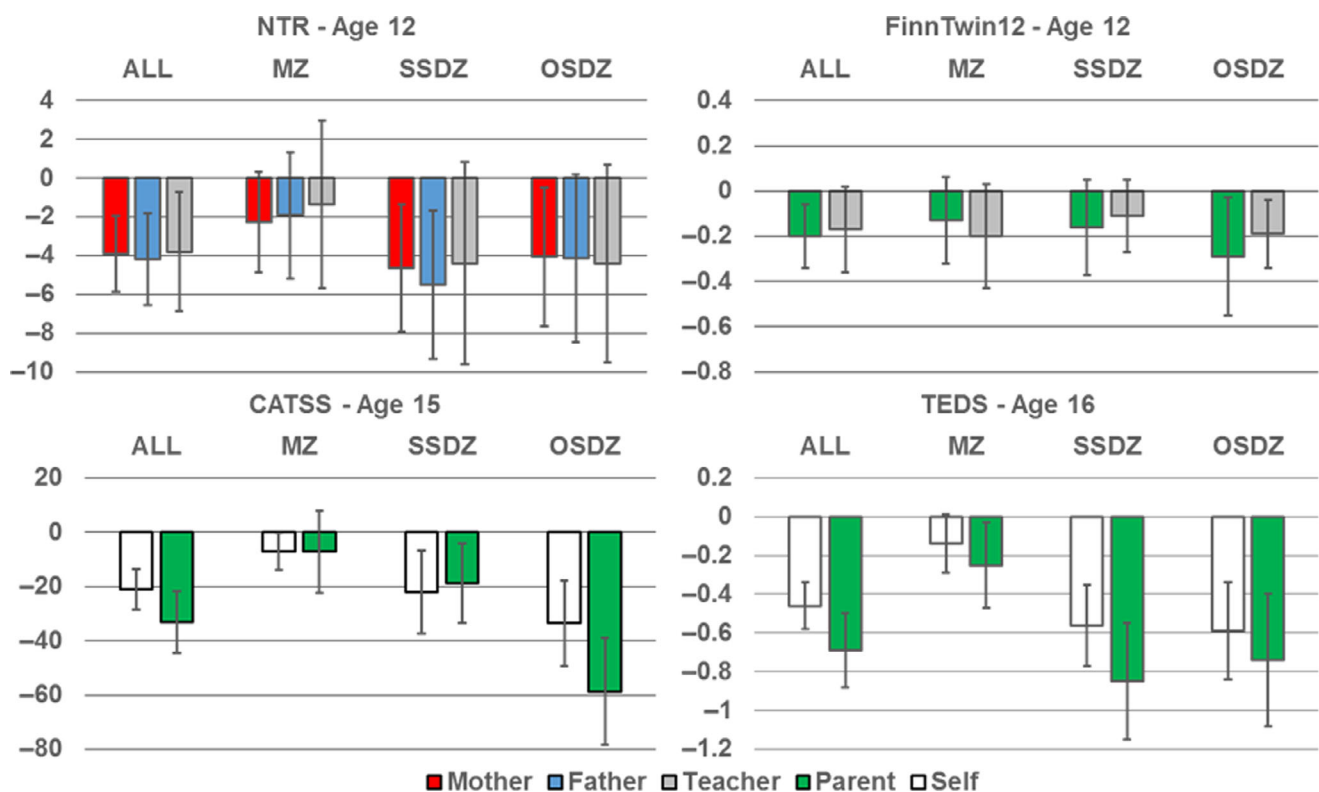


Figure 4 Within-family analyses on the association between aggression and academic performance. Unstandardized regression coefficient (with 95% confidence intervals) indicate the effect of one unit increase in the aggression measure on academic performance in a model that controlled for age and sex. NTR, Netherlands Twin Registry; FinnTwin12, Finnish Twin Cohort study; CATSS, Child and Adolescent Twin Study in Sweden; TEDS, Twins Early Development Study. FinnTwin12 and CATSS had teacher-rated grade point average as dependent variables and NTR and TEDS had standardized test scores as dependent variables. See Table S10 for additional models in NTR (age 7 and 10) and FinnTwin12 (age 14) [Colour figure can be viewed at wileyonlinelibrary.com]

Between-family analyses indicated a robust modest negative association between aggression and academic performance. This relationship is not an artefact arising from rater bias, for example when the same teachers rate children's aggression and school grades, nor was it age-dependent. This negative association was evident consistently in childhood/adolescence across ages when children have compulsory schooling in these four countries. Finding similar associations when using teacher-rated GPA and standardized test scores suggests that the associations between aggression and academic performance are not dependent on the type of performance rating. Looking across cohorts, the correlations between aggression and academic performance were strongest in adolescence at ages 15 and 16 which corresponds to the findings that adolescent aggression is more strongly associated with negative outcomes than childhood aggression (Pulkkinen, 2018). Considering analyses with parental and self-rated aggression, we observed significant heterogeneity between studies in individual level analyses. And in both cases, age 15 and 16 correlations with academic performance were greater than at age 12. However, we note that it was not possible to test definitely if the aggression-academic performance association was stronger at older ages because different aggression instruments were used in different cohorts that in turn represented different ages.

Within-family analyses of twin pairs showed that the associations between aggression and academic performance are evident even when comparing within-twin-pair differences in aggression in relation to within-twin-pair differences in academic performance. When including all twin pairs, the negative association between aggression and academic performance was statistically significant in 14/17 analyses. Twins share common environmental influences coming from having the same home (e.g. parental socioeconomic status and parental mental problems) and generally the same school and share genetic background either as full siblings (DZ pairs) or completely (as MZ pairs). These analyses controlled also for shared genetic effects, partly in DZs and fully in MZs. In within-family analyses, the sample size was smaller when including only MZ, SSDZ or OSDZ pairs, but meta-analytical approach separately analysing parental and self-ratings of aggression showed that correlations with academic performance were statistically significant, even in MZ pairs. The pattern of our within-twin-pair analyses indicated partial genetic confounding, that is only modestly smaller within-family correlations compared to between-family correlations and MZ within-twin-pair correlations about half as large as SSDZ within-twin-pair correlations. However, it should be noted that there was significant heterogeneity in within-family correlations of DZ twin pairs whereas no heterogeneity was observed in correlations of MZ twin pairs.

One advantage of within-twin analyses in same-sex pairs is the ability to fully control for sex. Considering shared environmental effects, twins from same families share their parents' educational level and socioeconomic status, factors that are related to both offspring aggression and academic performance (Assari, Caldwell & Bazargan, 2019). Our within-twin-pair analyses indicated that shared environmental effects attenuated the correlation between aggression and academic performance only to a small extent suggesting that shared environmental factors such as parental education or socioeconomic background do not explain this association. However, it should be noted that many factors in home environment—even socioeconomic status—are also affected by genetic effects. Within-family comparisons in MZ twins controlled also fully for genetic effects and these analyses strongly indicated shared genetic effects underlying the link between aggression and academic performance. Future studies should investigate if polygenic score of educational attainment is associated with aggression or if polygenic scores of aggression explain individual differences in academic performance. Eventually, studies using polygenic scores could use Mendelian randomization to test the direction of causality: does higher aggression cause poorer academic performance or is it the other way around?

One of the limitations of our study was the inclusion of only one instrument per country to assess aggression making it impossible to distinguish between instrument and country effects. In Sweden and the UK, the SDQ was administered and these two studies indicated very similar pattern of results in spite of using different types of academic performance measures (teacher-rated GPA in Sweden and standardized test score in the UK). Data with two aggression instruments in the same sample would have allowed us to investigate if the associations with different instrument are similar or different in magnitude. Nevertheless, our results were consistent across countries, raters and age, indicating robustness of the aggression-academic performance association.

Another limitation is that we did not include early childhood cognitive ability measures. Such data would have permitted a test of whether early cognitive ability predicts academic performance in adolescence and if aggression has any additive or mediating effect on these associations. There is evidence for a pathway from higher aggression to lower academic performance in childhood and adolescence (Van der Ende, Verhulst, & Tiemeier, 2016), with childhood aggression predicting academic performance in adolescence and educational attainment in adulthood even after controlling for intelligence (Dubow, Huesmann, Boxer, Pulkkinen, & Kokko, 2006; Masten et al., 2005). However, the nature of the association is unclear; the reverse causal

pathway may also be present, as a low intelligence was predictive of higher antisocial behaviour in childhood (Koenen et al., 2006). A third scenario involves confounding, because childhood intelligence predicted criminality, poorer mental health and substance abuse problems in early adulthood, but the associations were much weaker after controlling for children's behavioural problems and family background (Fergusson, Horwood, & Ridder, 2005a).

We note that also personality could explain, possible via genetic effects, a robust aggression–academic performance association. Aggression is closely linked with self-control (Denson, DeWall, & Finkel, 2012)—which is also highly heritable (Willems, Boesen, Li, Finkenauer & Bartels, 2019)—that predicts academic performance even over and above intelligence (Duckworth & Seligman, 2005). Childhood self-control indicated by constructive/prosocial behaviour (active coping with a problem, positive thinking and consideration of others with helpfulness and empathy) is associated with school success, low levels of dropping out of education, career orientation, occupational status and income (Pulkkinen, 2017, p. 263–266). In fact, prosocial behaviour has found to be even stronger predictor of academic performance than aggression (Caprara, Barbaranelli, Pastorelli, Bandura, & Zimbardo, 2000).

Strengths of our study include having four large samples of twins from longitudinal studies. Our comprehensive study included aggression assessments from different raters and at different ages and we used both teacher-rated GPA and standardized test scores. Within-family analyses allowed us to test aggression–academic performance relationships by controlling for unmeasured shared genetic and environmental influences.

In conclusion, our results indicate a robust negative association between aggression and academic performance across countries and across childhood into adolescence. Two main findings from twin analyses were that shared environmental and genetic effects are important for the association between aggression and academic performance across childhood and adolescence but even when controlling for genetic effects there is still some evidence for a negative association between aggression and academic performance.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article:

Table S1. Items of aggression scales.

Table S2. Sex differences in aggression.

Table S3. Sex differences in academic performance.

Table S4. Between-family regression models in the Netherlands Twin Registry with age 7 aggression measures.

Table S5. Between-family regression models in the Netherlands Twin Registry with age 10 aggression measures.

Table S6. Between-family regression models in the Netherlands Twin Registry with age 12 aggression measures.

Table S7. Between-family regression models in the FinnTwin12.

Table S8. Between-family regression models in the Child and Adolescent Twins Study in Sweden.

Table S9. Between-family regression models in the Twins Early Development Study.

Table S10. Correlations between aggression and academic performance by different raters and sex in different twin cohorts.

Table S11. Within-family regression models.

Table S12. Within twin pair correlations between aggression and academic performance by different raters in different twin cohorts.

Figure S1. Meta-analysis of within-family correlations between parental rated aggression and academic performance including MZ twin pairs.

Figure S2. Meta-analysis of within-family correlations between self-rated aggression and academic performance including MZ twin pairs.

Figure S3. Meta-analysis of within-family correlations between parental rated aggression and academic performance including DZ same-sex twin pairs.

Figure S4. Meta-analysis of within-family correlations between self-rated aggression and academic performance including DZ same-sex twin pairs.

Figure S5. Meta-analysis of within-family correlations between parental rated aggression and academic performance including DZ opposite-sex twin pairs.

Figure S6. Meta-analysis of within-family correlations between self-rated aggression and academic performance including DZ opposite-sex twin pairs.

Acknowledgements

This work is funded by the European Union 7th framework (FP7/2003–2013) through “Aggression in Children: Unraveling gene-environment interplay to inform Treatment and InterventiON strategies” (ACTION) consortium under grant agreement no 602768. FinnTwin12 data collection was supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (grants 213506, 129680), the Academy of Finland (grants 100499, 205585, 118555, 308248 and 312073 to J.K.), National Institute of Alcohol Abuse and Alcoholism (grants AA-12502, AA-00145, and AA-09203 to R.J.R.). J.K. has been supported as Academy professor by the Academy of Finland (265240, 263278). E.V. is supported as Academy research fellow by the Academy of Finland (314639). The Child and Adolescent Twin Study in Sweden study is supported by the Swedish Council for Working Life, the Söderström-Königska Foundation, and the Swedish Research Council (Medicine, Humanities and Social Science, grant number 2017-02552 and SIMSAM). NTR acknowledges the following support: Longitudinal data collection from teachers of Dutch twins and their siblings (NWO-481-08-011); Innovative learning materials and methods (Netherlands Initiative Brain and

Cognition NWO-056-31-001); Netherlands Twin Registry Repository (NWO 480-15-001/674); Consortium on Individual Development, which is funded through the Gravitation program of the Dutch Ministry of Education, Culture and Science and the Netherlands Organization for Scientific Research (NWO 0240-001-003) and the Royal Netherlands Academy of Science Professor Award (PAH/6635) to D.I.B.. TEDS is supported by a program grant to R.P. from the UK Medical Research Council (MR/M021475/1), with additional support from the US National Institutes of Health (AG046938). R.P. is supported by a Medical Research Council Professorship award (G19/2). K.R. is supported by the

Sir Henry Wellcome Postdoctoral Fellowship. The authors thank the participating twins, their parents and teachers for their contribution. The authors have declared that they have no competing or potential conflicts of interest.

Correspondence

Eero Vuoksimaa, Institute for Molecular Medicine Finland, P.O. Box 20 (Tukholmankatu 8), 00014 University of Helsinki, Helsinki, Finland; Email: eero.vuoksimaa@helsinki.fi

Key points

- There is a negative relationship between aggression and academic performance, but this association is not well understood.
- We performed a comprehensive assessment of the aggression–academic performance relationship in >27,000 children and adolescents from four European twin cohorts.
- Results indicated a robust negative association between aggression and academic performance that is evident across childhood development (ages 7–16), raters (parents, teacher, self) and different instruments.
- Aggression–academic performance relationship was also evident in within-twin-pair analyses but weaker in monozygotic pairs indicating partial genetic confounding.
- The relationship between aggression and academic performance is evident across compulsory education and is not an artefact arising from rater bias when the same teachers rate children's aggression and school grades.

References

- Achenbach, T.M., Ivanova, M.Y., & Rescorla, L.A. (2017). Empirically based assessment and taxonomy of psychopathology for ages 1(1/2)-90+ years: Developmental, multi-informant, and multicultural findings. *Comprehensive Psychiatry*, 79, 4–18.
- Anckarsater, H., Lundstrom, S., Kollberg, L., Kerekes, N., Palm, C., Carlstrom, E., ... & Lichtenstein, P. (2011). The child and adolescent twin study in Sweden (CATSS). *Twin Research and Human Genetics*, 14, 495–508.
- Anderson, C.A., & Bushman, B.J. (2002). Human aggression. *Annual Review of Psychology*, 53, 27–51.
- Assari, S., Caldwell, C.H., & Bazargan, M. (2019). Association between parental educational attainment and youth outcomes and role of race/ethnicity. *JAMA Network Open*, 2, e1916018.
- Bartels, M., Hendriks, A., Mauri, M., Krapohl, E., Whipp, A., Bolhuis, K., ... & Boomsma, D.I. (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, 1105–1121.
- Basisonderwijs, Eindtoets (2002). *Eindtoets basisonderwijs*. Arnhem: Citogroep.
- Boomsma, D.I., Vink, J.M., van Beijsterveldt, T.C., de Geus, E.J., Beem, A.L., Mulder, E.J., ... & van Baal, G.C.M. (2002). Netherlands twin register: A focus on longitudinal research. *Twin Research*, 5, 401–406.
- Bynner, J., & Parsons, S. (2002). Social exclusion and the transition from school to work: The case of young people not in education, employment, or training (NEET). *Journal of Vocational Behavior*, 60, 289–309.
- Caprara, G.V., Barbaranelli, C., Pastorelli, C., Bandura, A., & Zimbardo, P.G. (2000). Prosocial foundations of children's academic achievement. *Psychological Science*, 11, 302–306.
- De Moor, M.H., Boomsma, D.I., Stubbe, J.H., Willemsen, G., & de Geus, E.J. (2008). Testing causality in the association between regular exercise and symptoms of anxiety and depression. *Archives of General Psychiatry*, 65, 897–905.
- de Zeeuw, E.L., van Beijsterveldt, C.E., Glasner, T.J., de Geus, E.J., & Boomsma, D.I. (2016). Arithmetic, reading and writing performance has a strong genetic component: A study in primary school children. *Learning and Individual Differences*, 47, 156–166.
- Denson, T.F., DeWall, C.N., & Finkel, E.J. (2012). Self-control and aggression. *Current Directions in Psychological Science*, 21, 20–25. <http://doi.org/10.1177/0963721411429451>
- Dick, D.M., Johnson, J.K., Viken, R.J., & Rose, R.J. (2000). Testing between-family associations in within-family comparisons. *Psychological Science*, 11, 409–413.
- Dubow, E.F., Huesmann, L.R., Boxer, P., Pulkkinen, L., & Kokko, K. (2006). Middle childhood and adolescent contextual and personal predictors of adult educational and occupational outcomes: A mediational model in two countries. *Developmental Psychology*, 42, 937–949.
- Duckworth, A.L., & Seligman, M.E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, 16, 939–944.
- Fergusson, D.M., Horwood, L.J., & Ridder, E.M. (2005a). Show me the child at seven II: Childhood intelligence and later outcomes in adolescence and young adulthood. *Journal of Child Psychology and Psychiatry*, 46, 850–858.
- Fergusson, D.M., Horwood, L.J., & Ridder, E.M. (2005b). Show me the child at seven: The consequences of conduct problems in childhood for psychosocial functioning in adulthood. *Journal of Child Psychology and Psychiatry*, 46, 837–849.
- Haworth, C.M., Davis, O.S., & Plomin, R. (2013). Twins early development study (TEDS): A genetically sensitive investigation of cognitive and behavioral development from

- childhood to young adulthood. *Twin Research and Human Genetics*, 16, 117–125.
- Haworth, C.M., Wright, M.J., Luciano, M., Martin, N.G., de Geus, E.J., van Beijsterveldt, C.E., ... & Plomin, R. (2010). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular Psychiatry*, 15, 1112–1120.
- Hicks, B.M., Johnson, W., Iacono, W.G., & McGue, M. (2008). Moderating effects of personality on the genetic and environmental influences of school grades helps to explain sex differences in scholastic achievement. *European Journal of Personality*, 22, 247–268.
- Hinshaw, S.P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin*, 111, 127–155.
- Hudziak, J.J., van Beijsterveldt, C.E., Bartels, M., Rietveld, M.J., Rettew, D.C., Derks, E.M., & Boomsma, D.I. (2003). Individual differences in aggression: Genetic analyses by age, gender, and informant in 3-, 7-, and 10-year-old Dutch twins. *Behavior Genetics*, 33, 575–589.
- Johnson, W., McGue, M., & Iacono, W.G. (2009). School performance and genetic and environmental variance in antisocial behavior at the transition from adolescence to adulthood. *Developmental Psychology*, 45, 973–987.
- Kaprio, J., Pulkkinen, L., & Rose, R.J. (2002). Genetic and environmental factors in health-related behaviors: Studies on Finnish twins and twin families. *Twin Research*, 5, 366–371.
- Koenen, K.C., Caspi, A., Moffitt, T.E., Rijdsdijk, F., & Taylor, A. (2006). Genetic influences on the overlap between low IQ and antisocial behavior in young children. *Journal of Abnormal Psychology*, 115, 787–797.
- Kokko, K., & Pulkkinen, L. (2000). Aggression in childhood and long-term unemployment in adulthood: A cycle of maladaptation and some protective factors. *Developmental Psychology*, 36, 463–472.
- Krapohl, E., Rimfeld, K., Shakeshaft, N.G., Trzaskowski, M., McMillan, A., Pingault, J.B., ... & Plomin, R. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 15273–15278.
- Latvala, A., Rose, R.J., Pulkkinen, L., Dick, D.M., Korhonen, T., & Kaprio, J. (2014). Drinking, smoking, and educational achievement: Cross-lagged associations from adolescence to adulthood. *Drug and Alcohol Dependence*, 137, 106–113.
- Lewis, G.J., Asbury, K., & Plomin, R. (2017). Externalizing problems in childhood and adolescence predict subsequent educational achievement but for different genetic and environmental reasons. *Journal of Child Psychology and Psychiatry*, 58, 292–304.
- Masten, A.S., Roisman, G.I., Long, J.D., Burt, K.B., Obradovic, J., Riley, J.R., ... Tellegen, A. (2005). Developmental cascades: Linking academic achievement and externalizing and internalizing symptoms over 20 years. *Developmental Psychology*, 41, 733–746.
- McGue, M., Osler, M., & Christensen, K. (2010). Causal inference and observational research: The utility of twins. *Perspectives on Psychological Science*, 5, 546–556.
- Oliver, B.R., & Plomin, R. (2007). Twins' early development study (TEDS): A multivariate, longitudinal genetic investigation of language, cognition and behavior problems from childhood through adolescence. *Twin Research and Human Genetics*, 10(1), 96–105.
- Porsch, R.M., Middeldorp, C.M., Cherny, S.S., Krapohl, E., van Beijsterveldt, C.E., Loukola, A., ... Bartels, M. (2016). Longitudinal heritability of childhood aggression. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, 171, 697–707.
- Pulkkinen, L. (2017). *Human development from middle childhood to middle adulthood. Growing up to be middle-aged (in collaboration with Katja Kokko)*. London: Routledge.
- Pulkkinen, L. (2018). Longitudinal study of personality and social development; Insights about aggression after five decades. In A.T. Vazsonyi, D.J. Flannery, & M. DeLisi (Eds.), *The Cambridge handbook of violent behavior and aggression* (2nd ed., pp. 31–51). Cambridge: Cambridge University Press.
- Pulkkinen, L., Kaprio, J., & Rose, R.J. (1999). Peers, teachers and parents as assessors of the behavioural and emotional problems of twins and their adjustment: The Multidimensional Peer Nomination Inventory. *Twin Research*, 2, 274–285.
- Rimfeld, K., Kovas, Y., Dale, P.S., & Plomin, R. (2015). Pleiotropy across academic subjects at the end of compulsory education. *Scientific Reports*, 5, 11713.
- Rimfeld, K., Malanchini, M., Spargo, T., Spickernell, G., Selzam, S., McMillan, A., ... & Plomin, R. (2019). Twins Early Development Study: A genetically sensitive investigation into behavioral and cognitive development from infancy to emerging adulthood. *Twin Research and Human Genetics*, 22, 508–513.
- Schwarzer, G. (2007). meta: An R package for meta-analysis. *R News*, 7, 40–45.
- Shakeshaft, N.G., Trzaskowski, M., McMillan, A., Rimfeld, K., Krapohl, E., Haworth, C.M., ... Plomin, R. (2013). Strong genetic influence on a UK nationwide test of educational achievement at the end of compulsory education at age 16. *PLoS One*, 8, e80341.
- Tremblay, R.E., Vitaro, F., & Cote, S.M. (2018). Developmental origins of chronic physical aggression: A bio-psycho-social model for the next generation of preventive interventions. *Annual Review of Psychology*, 69, 383–407.
- van Beijsterveldt, C.E., Groen-Blokhuis, M., Hottenga, J.J., Franic, S., Hudziak, J.J., Lamb, D., ... Boomsma, D.I. (2013). The young Netherlands twin register (YNTR): Longitudinal twin and family studies in over 70,000 children. *Twin Research and Human Genetics*, 16, 252–267.
- Van der Ende, J., Verhulst, F.C., & Tiemeier, H. (2016). The bidirectional pathways between internalizing and externalizing problems and academic performance from 6 to 18 years. *Development and Psychopathology*, 28, 855–867.
- Voyer, D., & Voyer, S.D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140, 1174–1204.
- Willems, Y.E., Boesen, N., Li, J., Finkenauer, C., & Bartels, M. (2019). The heritability of self-control: A meta-analysis. *Neuroscience and Biobehavioral Reviews*, 100, 324–334.

Accepted for publication: 4 May 2020