

University of Groningen

Benefits of Bullying? A Test of the Evolutionary Hypothesis in Three Cohorts

Kretschmer, Tina; la Roi, Chaïm; van der Ploeg, Rozemarijn; Veenstra, René

Published in:
Journal of Research on Adolescence

DOI:
[10.1111/jora.12675](https://doi.org/10.1111/jora.12675)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2022

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Kretschmer, T., la Roi, C., van der Ploeg, R., & Veenstra, R. (2022). Benefits of Bullying? A Test of the Evolutionary Hypothesis in Three Cohorts. *Journal of Research on Adolescence*, 32(3), 1178-1193.
<https://doi.org/10.1111/jora.12675>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Benefits of Bullying? A Test of the Evolutionary Hypothesis in Three Cohorts

Tina Kretschmer[†] 
University of Groningen

Chaim la Roi[†]
Institute for Future Studies & Swedish Institute for Social
Research

Rozemarijn van der Ploeg, and René Veenstra
University of Groningen

Recent work on bullying perpetration includes the hypothesis that bullying carries an evolutionary advantage for perpetrators in terms of health and reproductive success. We tested this hypothesis in the National Child Development Study ($n = 4998$ male, $n = 4831$ female), British Cohort Study 1970 ($n = 4261$ male, $n = 4432$ female), and TRacking Adolescents' Individual Lives Survey ($n = 486$ male, $n = 521$ female), where bullying was assessed in adolescence (NCDS, BCS70: age 16, TRAILS: age 14) and outcomes in adulthood. Partial support for the evolutionary hypothesis was found as bullies had more children in NCDS and engaged in sexual intercourse earlier in TRAILS. In contrast, bullies reported worse health in NCDS and BCS70.

Key words: bullying perpetration – evolutionary hypothesis – longitudinal cohort study

The literature on bullying victimization has surged in the past decade, with studies usually concluding that victims of bullying pay a high price in terms

of mental and physical health (Schoeler, Duncan, Cecil, Ploubidis, & Pingault, 2018). Negative outcomes are often implied for bullying perpetrators as well and reviews suggest modest to moderate effect sizes for problem outcomes in bullies (Ttofi, Farrington, Lösel, Crago, & Theodorakis, 2016; Vrijen, Wiertsema, Ackermans, van der Ploeg, & Kretschmer, 2021). However, negative outcomes of bullying perpetration are not universal: After controlling for childhood risk, bullies mastered normative developmental tasks just as well as nonbullies (Kretschmer et al., 2018) and did not fare worse on psychological adjustment, health, or wealth (Copeland, Wolke, Angold, & Costello, 2013; Wolke, Copeland, Angold, & Costello, 2013). Some suggest that bullying has an adaptive function and is instrumental to greater health and reproductive fitness (Koh & Wong, 2017; Volk, Camilleri, Dane, & Marini, 2012; Volk, Dane, Dane, Marini, & Vaillancourt, 2015; Volk, Della Cioppa, Della Cioppa, Earle, & Farrell, 2015). Here, we provide an extensive empirical test of the hypothesis that bullying might carry an evolutionary advantage for perpetrators, utilizing data from three cohorts, of which two long-term: The National Child Development Study (NCDS) has followed participants until age

[†]These authors have contributed to the manuscript in equal parts and share first authorship.

We are grateful to the Centre for Longitudinal Studies (CLS), UCL Institute of Education, for the use of the NCDS and BCS70 data, and to the UK Data Service for making them available. Neither CLS nor the UK Data Service bear any responsibility for the analysis or interpretation of these data. This research is also part of the TRacking Adolescents' Individual Lives Survey (TRAILS). Participating centers of TRAILS include various departments of the University Medical Center and University of Groningen, the University of Utrecht, the Radboud Medical Center Nijmegen, and the Parnassia Group, all in the Netherlands. TRAILS has been financially supported by various grants from the Netherlands Organization for Scientific Research (NWO), ZonMW, GB-MaGW, the Dutch Ministry of Justice, the European Science Foundation, the European Research Council, BBMRI-NL, and the participating universities. We are grateful to everyone who participated in this research or worked on this project. Preparation of the manuscript has been supported by a European Research Council (ERC) Starting Grant awarded to Tina Kretschmer under the Horizon 2020 Research and Innovation program (Grant Agreement Number 757364, Title: Ghosts from the Past-Consequences of Adolescent Peer Relations Across Contexts and Generations). Chaim la Roi acknowledges funding from two research projects funded by the Swedish Research Council for Health, Working Life and Welfare (FORTE) when preparing this manuscript for submission: YOUNG [2012-1741] and Interlocking inequalities [2016-07099].

Requests for reprints should be sent to Tina Kretschmer, Department of Pedagogy and Educational Sciences, Faculty of Behavioral and Social Sciences, University of Groningen, Grote Rozenstraat 38, Groningen 9712TJ, The Netherlands. E-mail: t.kretschmer@rug.nl

© 2021 The Authors. *Journal of Research on Adolescence* published by Wiley Periodicals LLC on behalf of Society for Research on Adolescence

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

DOI: 10.1111/jora.12675

55, the British Cohort Study (BCS70) has followed participants until age 42, and the Tracking Adolescents' Individual Lives Survey (TRAILS) has followed participants until age 26. All include information on bullying perpetration in adolescence and health and reproductive outcomes several years later.

Evolutionary Perspective on Bullying

From an evolutionary standpoint, humans have evolved to display behaviors that increase their likelihood to survive and reproduce. If bullying perpetration is advantageous, the behavior itself or a trait for which it is a proxy should thus contribute to survival and reproduction. An evolutionary perspective on bullying would thus suggest that the behavior is instrumental to forming a "pecking order", similar to other species in which a group's hierarchy is relatively stable. Established hierarchies prevent constant power battles and facilitate group stability (Zwaan, Dijkstra, & Veenstra, 2013). Bullies are thought to demonstrate strength and their dominance might prevent them from becoming targets themselves (Archer & Benson, 2008). In other words, adolescents are less likely to pick fights with those at greater power which could carry health benefits such as lower stress levels than among those lower in the hierarchy. Indeed, bullies show the lowest increase of C-reactive protein, a marker of systematic inflammation, between childhood and early adulthood, suggesting better somatic health years later in bullies compared with victims and noninvolved peers (Copeland et al., 2014). Better health, in turn, is important for survival, suggesting that bullies might be at an advantage. However, it is not clear whether such health benefits extend over time as a longitudinal study did not imply health benefits for bullies (Matthews, Jennings, Lee, & Pardini, 2017).

With respect to reproductive outcomes, it is believed that male bullies are more dominant, display physical strength, and attain more material resources and female bullies are more attractive—traits that might contribute to a higher social status and greater opportunities for mating and reproduction (Volk et al., 2012). Indeed, conform to a frequently high position in the social hierarchy, bullies are more likely to date, have sex, and report a greater number of partners in adolescence and early adulthood (Dane, Marini, Volk, & Vaillancourt, 2017; Farrell & Vaillancourt, 2019; Provenzano, Dane, Farrell, Marini, & Volk, 2018; Volk, Dane, et al., 2015). These studies suggest that bullies appear to be at an advantage in an evolutionary

sense as they are sexually attractive to a greater number of potential partners, which should increase their reproductive opportunities. Although bullies start dating earlier, their relationships are less supportive and less durable (Connolly, Pepler, Craig, & Taradash, 2000), possibly paving the way for an early onset of changing relationship partners, which, at least for men, can be seen as advantageous. That is, a greater number of partners again increase the number of reproductive opportunities and by extension the number of offspring. Taken together, bullying might be beneficial through averting stress-related somatic problems and simplifying the processes of dating and mating but systematic longitudinal research into adult outcomes of bullying that would point at adaptation is lacking.

Importantly, evolutionary benefits might vary across bullying perpetrators: They might not be available to bully victims who tend to take the lowest rank in the peer hierarchy (Koh & Wong, 2017), are not favored as dating partners (Dane et al., 2017), and often have bad health (Wolke et al., 2013). Their bullying is considered retaliatory rather than proactive, which contrasts evolutionary assumptions about "pure" bullies. Thus, whether or not a bully is simultaneously victimized might moderate associations between bullying perpetration and outcomes. Similarly, although bullying is widely recognized as a plight, "pure" bullies but not bully victims are generally high in social standing among their peers (Sentse, Kretschmer, & Salmivalli, 2015; van der Ploeg et al., 2020), which, in turn, is linked to better health (Garza, Glenn, Mistry, Ponce, & Zimmerman, 2017; Mundt & Zakletskaia, 2014; Sweeting & Hunt, 2014). Popular adolescents are also more successful on the dating market (Savickaitė, Dijkstra, Kreager, Ivanova, & Veenstra, 2020). As such, bullying should be associated with health and reproductive outcomes when coupled with high popularity.

Current Study

It has been suggested that bullying perpetration might be evolutionarily advantageous (Volk et al., 2012) through promoting health and opportunities for reproduction but rigorous longitudinal studies that test this assumption are scarce. This is a problem as we would expect to see a possible reproductive advantage of bullies reflected in a greater number of partners and offspring, which can most meaningfully be assessed once someone's reproductive period has been concluded. To be able to successfully pass on one's genes, people need to be

healthy, and the healthy status needs to be maintained for as long as possible to extend the period of potential reproduction and assist offspring in caring for grandchildren. As such, benefits of bullying can best be examined over the life course. To this end, we tested adult outcomes of adolescent bullying perpetration in three cohorts of different age. Based on Volk's evolutionary hypothesis of bullying, we expected that bullies would rate their health better than nonbullies, would engage in sexual activity earlier, have a greater number of partners, as well as have given birth to or fathered more children. Moreover, we expected that positive effects of bullying perpetration would be stronger for popular bullies compared with unpopular bullies and to be weaker for bully victims compared with pure bullies.

We used data from NCDS, where global health and reproductive success were assessed at age 55, BCS70, where global health and reproductive success were assessed at age 42, and TRAILS, where physical health and reproductive success were assessed up to age 26. Whereas the long follow-up in NCDS allows for studying links between bullying and lifetime reproductive success, BCS70 contains information on adolescents' popularity and victimization as moderators for whether or not bullying is likely to be associated with adaptive outcomes. TRAILS has assessed bullying, popularity, and victimization using peer nominations, whereas parent and teacher reports were used in NCDS and BCS70, and contains more detailed information on early adult sexual behavior.

Expression of bullying perpetration differs between girls and boys, and outcomes are not consistent across sex in studies with an evolutionary focus (Dane et al., 2017), we thus stratified analyses. We included similar covariates as used in previous studies on outcomes of bullying namely family SES, family disruption, and mental health (i.e., childhood psychopathology) (Copeland et al., 2013; Takizawa, Maughan, & Arseneault, 2014) as well as covariates with a potential effect on the outcomes of interest, namely health and BMI as precursors of adult health and pubertal timing as precursor of reproductive behavior. The study was preregistered on OSF; modifications are described in an amendment. Both documents can be found [here](#).

METHOD

Ethics approval was not needed for use of secondary data as determined by the Ethics Review

Committee of the institution at which this study was conducted (Faculty of Behavioral and Social Sciences, University of Groningen).

National Child Development Study

Sample. National Child Development Study began with over 17,000 children born in Britain in 1 week in 1958. Originally focusing on perinatal mortality and risk factors for stillbirth and neonatal death, NCDS retraced participants at various points throughout their lives. An overview of demographic data, assessments up to seventh follow-up (2003), and retention rates is provided elsewhere (Power & Elliott, 2006). Subsequent follow-ups are described by the Centre of Longitudinal Studies (<https://cls.ucl.ac.uk/cls-studies/1958-national-child-development-study/>). At age 55, ~52% of participants were still in the study. Table 1 shows data availability for each construct, and Table S1 depicts associations between predictor and covariates and nonresponse on outcome variables used in this study. Analyses were conducted on 4998 male and 4831 female participants.

Measures. *Bullying perpetration* was assessed from parents and teachers when participants were 16 years old as part of the 19-item Rutter Behaviour Scale (Rutter, Tizard, & Whitmore, 1970) ("Bullies other children"). We used the single item (1 = *does not apply*, 2 = *applies somewhat*, 3 = *certainly applies*) from both reporters and constructed a dichotomous indicator with children who were affirmed to be bullies at least somewhat by parents or teachers being assigned a score of 1 (11.6%), and the others a score of 0.

Adult health was assessed from participants at age 55, using the item: "In general, would you say your health is...?" with categories 1 = *excellent*, 2 = *very good*, 3 = *good*, 4 = *fair*, and 5 = *poor*. We used this outcome as continuous variable, $M = 3.35$, $SD = 1.07$.

Mating and reproductive success were assessed using the total number of partners with whom the participant had cohabited for 1 month or more (Hancock, 2017) and total number of children (biological children only) as reported between age 16 and age 50. Most participants had reported on one or two partners (74.4%). Note that this score refers to number of romantic partners. The number of children ranged from none (18.6%) to 14 ($n = 1$), with most participants being parent to one, two, or three children. To account for extreme outliers, we winsorized raw scores at 95%, referring to that we

TABLE 1
Descriptive Statistics and Mean Comparisons Between Bullying Perpetrators and Nonperpetrators

	n	Nonperpetrators M (SD)	Bullying Perpetrators M (SD)	t	p (t)
NCDS					
General health	5814	3.39 (1.05)	3.09 (1.14)	6.11	<.001
Number of partners	9778	1.18 (0.84)	1.15 (0.93)	1.30	.193
Number of children	5872	1.78 (1.13)	1.98 (1.17)	-4.00	<.001
Childhood SES	8029	3.39 (1.28)	2.88 (1.17)	11.49	<.001
Biological parents	9826	0.85 (0.36)	0.77 (0.42)	6.71	<.001
Family instability – parents divorced	9829	0.06 (0.23)	0.10 (0.30)	-5.20	<.001
Family instability – one or both parents dead	9829	0.05 (0.22)	0.06 (0.24)	-1.48	.139
Child psychopathology	9648	0.03 (0.18)	0.09 (0.29)	-9.18	<.001
Child health	8451	1.42 (1.51)	1.42 (1.56)	0.04	.966
Child BMI	8333	20.50 (2.39)	20.72 (2.43)	-2.63	.009
Pubertal development – age at menarche (girls)	3466	12.81 (1.35)	13.05 (1.52)	-2.83	.005
Pubertal development – pubic hair and breast (girls)	4153	3.12 (0.55)	3.13 (0.55)	-0.25	.800
Pubertal development – voice broken (boys)	3971	0.92 (0.27)	0.90 (0.29)	1.41	.158
Pubertal development – pubic hair (boys)	4286	2.62 (0.65)	2.60 (0.67)	0.65	.513
BCS70					
General health	5925	3.68 (1.05)	3.31 (1.19)	6.44	<.001
Number of partners	8693	1.22 (0.83)	1.19 (0.97)	1.07	.285
Number of children	5943	1.61 (1.13)	1.71 (1.22)	-1.76	.079
Popularity	4330	0.28 (0.45)	0.37 (0.48)	-2.58	.010
Victimization	3273	0.13 (0.26)	0.19 (0.31)	-2.94	.003
Childhood SES	5602	3.70 (1.35)	3.33 (1.50)	4.76	<.001
Biological parents	8347	0.81 (0.39)	0.71 (0.45)	5.99	<.001
Family instability – parents divorced	8355	0.14 (0.34)	0.25 (0.43)	-7.44	<.001
Family instability – one or both parents dead	7585	0.04 (0.18)	0.04 (0.20)	-0.76	.447
Child psychopathology	5209	0.05 (0.21)	0.09 (0.29)	-3.80	<.001
Child health	6143	1.07 (1.66)	1.19 (1.74)	-1.35	.176
Child BMI	5235	21.13 (2.62)	21.40 (2.79)	-1.83	.068
Pubertal development – age at menarche	4251	3.67 (1.30)	3.96 (1.34)	-3.44	<.001
TRAILS					
General health	631	3.19 (0.80)	3.17 (0.79)	0.20	.838
Physical complaints	637	0.29 (0.28)	0.29 (0.29)	0.10	.923
Number of partners	383	5.85 (2.76)	6.26 (3.07)	-1.14	.255
Number of children	845	0.07 (0.01)	0.08 (0.01)	-0.45	.654
Age at first sex	806	16.68 (2.67)	15.17 (2.30)	6.82	<.001
Popularity	1007	0.09 (0.12)	0.15 (0.15)	-6.27	<.001
Victimization	1007	0.01 (0.04)	0.05 (0.08)	-8.84	<.001
Childhood SES	990	0.14 (0.78)	-0.16 (0.81)	5.16	<.001
Family instability – parents divorced	933	0.16 (0.36)	0.25 (0.43)	-2.98	.003
Child psychopathology	914	0.16 (0.14)	0.22 (0.16)	-4.64	<.001
Child health	969	3.80 (0.76)	3.85 (0.78)	-0.77	.440
Child BMI	955	18.75 (3.06)	19.11 (3.47)	-1.49	.136
Pubertal development	726	3.22 (0.99)	2.83 (1.00)	4.27	<.001

corrected the 2.5% most extreme scores at either end to the nearest value. For right-skewed variables such as these, this amounts to correcting the 2.5% largest scores to the nearest value.

Covariates. For all covariates, assessments concurrent to (age 16) or preceding (age 7) the bullying perpetration assessment were selected. To the extent available, we used the same measures as in published studies. *Family SES* was measured using the age 11 assessment of father's occupational

social class (Wood et al., 2017), ranging from 1 = *unskilled* to 6 = *professional* ($M = 3.33$, $SD = 1.28$). *Family instability*, as reported by parents when the participants were 16, was measured with three dummy variables, identifying those who (1) lived with both biological parents (83.6%); (2) did not live with both biological parents due to divorce or separation of the parents (4%); (3) did not live with both biological parents due to death of (one of) the parents (3%). This strategy corresponds to

that used in earlier studies (Ely, Richards, Wadsworth, & Elliott, 1999). *Childhood psychopathology* was assessed using parent report at age 16, specifically the item “Has the study child ever been seen by a specialist for an emotional or behavioral problem?” which we dichotomized into 0 = *no* (95.7%) and 1 = *yes* (as an inpatient in a hospital, in a hospital outpatient department, at a child guidance clinic, elsewhere, 4.3%). We coded “don’t know” as missing. *Adolescent health* was assessed by a medical nurse, who, with the assistance of parent and study participant, completed a part of the age 16 assessment. We used the item “Has he/she attended a GP Surgery/Health Centre or been visited at home in the past 12 months?” with answer categories ranging from 1 = *no* to 6 = *yes, five or more times* ($M = 1.42$, $SD = 1.52$). The options “*Yes, don’t know frequency*” and “*Don’t know whether attended*” were coded as missing. *BMI* was assessed by dividing weight in kilograms by the square of height in meters, both derived from medical assessment at age 16 and included as a continuous variable ($M = 20.51$, $SD = 2.39$). Finally, *pubertal development* was assessed using medical questionnaire reports at age 16, with items for girls pertaining to age at menarche, breast development (*absent, intermediate, adult*), and axillary and pubic hair (*absent, sparse, intermediate, adult*) as examined by the nurse. For boys, whether or not breaking of the voice had already taken place (*yes/no*), and facial, axillary, and pubic hair development (*absent = 1, sparse = 2, intermediate = 3, adult = 4*) were assessed. Age at menarche ($M = 12.83$, $SD = 1.37$) and breaking of the voice (*yes = 92.2%*) were used as separate indicators, breast development (in girls), facial hair (in boys), and axillary as well as pubic hair development were combined to represent the mean score (girls: $M = 3.13$, $SD = 0.55$; boys: $M = 2.61$, $SD = 0.65$).

British Cohort Study (BCS70)

Sample. BCS70 began with over 17,000 children born in Britain in a single week in 1970. Originally focusing on social and biological characteristics of the mother in relation to neonatal morbidity, BCS70 retraced the original study participants at various points until age 46 years. An overview up to the age 34 follow-up (2004) including information on retention is published elsewhere (Elliott & Shepherd, 2006). Details on all follow-ups in BCS70 are available from the Centre of Longitudinal Studies (<https://cls.ucl.ac.uk/cls-studies/1970-british-cohort-study/>). At age 42, approximately 57% of

participants were still in the study. Table 1 shows data availability for each construct and Table S2 depicts links between predictor and covariates and nonresponse on outcomes used in this study. Analyses were conducted on 4261 male and 4432 female participants.

Measures. *Bullying perpetration* was assessed from parents when participants were 16 years old using a single item (“Teenager bullies others”), with answer categories: 1 = *certainly applies*, 2 = *applies somewhat*, 3 = *does not apply*. We constructed an indicator with individuals affirmed to bully others at least somewhat being assigned a score of 1 (7.2%), and the rest a score of 0.

Popularity was assessed using a single item (“I am popular”) in self-reports at age 16, with as answer categories: 1 = *applies very much*, 2 = *applies somewhat*, 3 = *does not apply*. We constructed a dummy with children who are affirmed to be very much popular being assigned a score of 1 (28.7%), and children who reported not being popular or being somewhat popular assigned a score of 0.

Victimization was assessed using three dichotomous items in self-reports at age 16 (“other pupils in school often fall out with me”, “often feel lonely at school”, “other pupils say nasty things about me”). We recoded the items such that 0 = *no* and 1 = *yes* (*don’t know* was coded as missing) and used the mean score ($M = 0.13$, $SD = 0.26$).

Adult health was assessed at the age 42 assessment, using the item: “In general, would you say your health is...?” with answer options 1 = *excellent*, 2 = *very good*, 3 = *good*, 4 = *fair*, and 5 = *poor*. We used the continuous indicator in analyses ($M = 3.60$, $SD = 1.07$).

Mating and reproductive success were derived from participants’ records of partners between age 16 and 42 (Hancock, 2016), including *Number of partners* ranging from 0 (29.6%) to 8 (0.1%), with most BCS70 participants reporting one, two, or three partners. Note that this score refers to number of partners with whom the participant has cohabited for 1 month or more. *Number of children* (biological children only) ranged from none (22.1%) to 10 ($n = 1$), with most participants (39%) being parent to two children at age 42. Raw scores were winsorized at 95%.

Covariates. For all covariates, assessments concurrent to (16-year) or preceding (5-year or 10-year) the bullying perpetration assessment were selected. To the extent available, measures were used in the same form as published before. *Family SES* was measured using the 10-year assessment of father’s

(or mother's, if father not present) occupational social class (Wood et al., 2017), ranging from 0 = *student* to 6 = *professional* ($M = 3.69$, $SD = 1.37$). *Family instability* was assessed using a similar strategy as in previous studies (Ely et al., 1999), such that three dummy variables were created of individuals who, as reported by parents, at age 16 lived with both biological parents (80%), did not live with both biological parents due to a divorce or separation of the parents (14.8%) and did not live with both biological parents due to the death of (one of) the parents (3.7%). *Childhood psychopathology* was assessed by a community medical officer as part of the age 16 medical examinations conducted at Child Health Clinics; we used the items "Is there any evidence that the study teenager has had any emotional or behavioral problems since age 10?" (*yes/no*) and "Is there evidence of any of the following psychological/psychiatric problems?" (*yes/no* to maladjustment/behavior disturbance, depression, aggression, appetite problems, psychosis, neurosis, suicide attempts/threats). We created a composite score with participants for whom the medical officer affirmed either of the two questions being assigned a score of 1 (5.25%), all other participants were assigned a score of 0. *BMI* was assessed by dividing weight in kilograms by the square of height in meters, both derived from medical assessment at age 16 and analyzed as a continuous variable. Raw scores were winsorized at 95% ($M = 21.15$, $SD = 2.64$). *Child health* was assessed in self-reports at age 16. We used the global item "How many times have you seen a doctor in the past two years?", with answer categories ranging from *never* = 0 to *six or more times* = 6. We used the item as a continuous variable ($M = 1.06$, $SD = 1.67$). Finally, *pubertal development* was assessed from parents at the age 16 assessment, however, for girls only. We used age at menarche (*before 11th birthday* = 1, *aged 11*, *aged 12*, *aged 13*, *aged 14*, *aged 15 or more*, *not yet commenced* = 7) and coded "commenced but don't know when" as missing ($M = 3.70$, $SD = 1.30$). The item was coded missing for boys as no information was available on their pubertal development at age 16 from either reporter.

TRacking Adolescents' Individual Lives Survey

Sample. TRacking Adolescents' Individual Lives Survey is a prospective cohort study of Dutch adolescents that started in 2001, with assessments at ages 11, 14, 16, 19, 22, and 26. In contrast to NCDS and BCS70, TRAILS did not start at birth

but when participants were in preadolescence and is a local sample. Details about TRAILS have been published in several reports (Oldehinkel et al., 2015). In brief, a total of 2,935 children were invited to participate of whom 2,229 (51% female) did so at T1. Retention was excellent with 96% at T2, 81% at T3, 84% at T4, 80% at T5, and 73% for T6. Table 1 shows data availability for each construct. Table S3 shows attrition analyses for the data used in this study. Analyses were conducted on 486 male and 521 female participants.

Measures. *Bullying perpetration.* At the first two assessments, a subsample of TRAILS participants took part in classroom-based assessments, in which TRAILS participants and their classmates nominated each other on a range of domains including bullying perpetration. At the age 14 assessment (used here), peer nominations were conducted in classrooms with at least three regular TRAILS participants or two participants on the condition that both had also participated in the earlier peer nominations assessment (Dijkstra, Cillessen, Lindenberg, & Veenstra, 2009). Peer nominations were assessed classroom-based, including non-TRAILS participants, with assessments taking place during regular lessons and lasting for about 15 min. After brief instructions in which a TRAILS staff member emphasized that information would be kept confidential, adolescents received the questionnaire with the names of the classmates listed. The teacher and TRAILS staff member remained in the classroom during the administration of the peer nominations.

Among other topics, adolescents were asked by whom they were bullied, for which they could nominate an unlimited number of classmates. The nominations received for bullying others were divided by the total number of participating pupils in the class, that is, the maximum number of nominations possible minus 1. This proportion score takes class size into account and ranges from 0 to 1, with higher scores indicating more bullying nominations. This procedure yields a continuous indicator but we dichotomized the score to compare bullies (any nomination, 24%) with nonbullies (no nomination), in line with NCDS and BCS70 and as required for the matching procedure applied in our analyses.

Popularity and victimization were assessed during the peer nomination procedure as well; that is, adolescents were also asked whom they bullied and indicated who is popular or someone they like to associate with. Proportion scores were created by dividing the number of nominations by the total

number of children in the class and dichotomized for subsequent analyses in order to maximize comparability with NCDS and BCS70. Victimization was dichotomized by distinguishing between respondents who had received no victimization nominations and those who had received at least one nomination (22%). Popularity was dichotomized by distinguishing between respondents with 25% highest proportion scores of nominations for being popular and the remaining 75% across the sample.

Adult health. At age 26, participants were asked how they evaluate their physical health in the past 30 days (1 = *bad* to 4 = *very good*) ($M = 3.18$, $SD = 0.82$). We additionally used the Adult Self Report (Achenbach & Rescorla, 2001) Physical Complaints scale ($\alpha = .78$), which consists of 12 items including “I feel tired without reason” and “I have stomach ache/racing heart/rashes or other skin problems without known medical reason” (0 = *does not apply*, 1 = *a little/sometimes*, 2 = *a lot/often*) as separate outcome ($M = 0.29$, $SD = 0.82$).

Mating and reproductive success. Several indicators of increased opportunities for mating and reproduction were assessed across the different waves and used to form the following constructs: *Age at first sex* was assessed from age 14 onward by asking whether participants ever had sexual sex and, if so, how old they were then ($M = 16.17$, $SD = 2.67$). From age 19 onwards, participants indicated with how many partners they had had sex in the past two years, the composite of those items provides an estimate of the *number of sexual partners* since age 17 ($M = 6.07$, $SD = 2.90$, range 0–15). Various measures were used to determine the *number of children* of TRAILS participants: At age 20, participants indicated whether they or their partner gave birth in the past 2 years. Using the Event History Calendar (Caspi et al., 1996), participants at age 23 also indicated whether they or their partner had been pregnant since age 16 and what the outcome of this (and additional) pregnancies was (child born, miscarriage, abortion). At ages 23 and 26, participants indicated whether they or their partner had given birth in the past two and a half years and the preceding five years. Given the low frequency of having had a child, a dummy was computed, distinguishing between “no children” (90%) and “any number of children”.

Covariates. *Family SES* was constructed from mothers’ and fathers’ educational and occupational levels and family income as measured at the first TRAILS wave (~age 11), thus preceding the bullying measurement. Educational level of parents was coded in five categories based on the International

Standard Classification of Occupations (Ganzeboom & Treiman, 1996). Disposable family income was measured on a scale ranging from less than €680 (1) to more than €3857 (9). Family SES was consequently operationalized as the average of the standardized five items ($\alpha = .84$); this indicator is commonly used in TRAILS analyses (Veenstra, Lindenberg, Oldehinkel, De Winter, & Ormel, 2006). *Parental divorce* prior to age 14 contrasted participants whose parents had separated to those whose biological parents were still together and was measured using information on whether the participant had experienced parental break-up as assessed from parents at age 11 and participants themselves at age 14 (22.4%). *Childhood psychopathology* was measured at the second TRAILS wave (~age 14) when bullying status was also assessed, using the Child Behavior Checklist Total Problems Scale (Achenbach & Rescorla, 2001), consisting of 119 items with very good reliability ($\alpha = .95$, $M = 0.18$, $SD = 0.15$). *Child health* was assessed from participants by asking “How would you rate your physical health during the past year” also at the second TRAILS wave (1 = *bad* to 5 = *very good*) ($M = 3.82$, $SD = 0.79$).

BMI was assessed by dividing weight in kilograms by the square of height in meters as recorded during the second TRAILS wave (~age 14) visit to the laboratory and corrected for clothing ($M = 19.00$, $SD = 3.21$). Finally, *pubertal development* was assessed at the second TRAILS wave using an adaptation of the Tanner stages scheme, where parents were shown five illustrations of male and female development and asked to indicate which illustration most fitted their son’s or daughter’s growth of pubic hair and penis/testes and breasts, respectively, coded 1 to 5, with higher scores indicating greater maturation ($M = 3.12$, $SD = 1.04$).

Analytic Strategy

For all samples, bullies and nonbullies were first compared on covariates. Second, missing data were imputed through multiple imputation by chained equations using ‘mice’ in R (Buuren & Groothuis-Oudshoorn, 2011). To improve imputation accuracy, we included auxiliary variables and created 50 imputed datasets. The Barnard-Rubin small sample adjustment was used for calculating *df* of test statistics (Barnard & Rubin, 1999). Moderators in BCS70 (social status, victimization) had missing values, which were imputed using the “transform, then impute” procedure (Von Hippel, 2009; White, Royston, & Wood, 2011). We included in this

procedure all participants with data on the bullying assessment.

Third, regression analyses were computed with health, number of partners, and number of children (all cohorts) as well as age at first sex (TRAILS) as outcomes. To control for confounding, we employed regression control in combination with Mahalanobis distance matching using the MatchingFrontier R package (King, Lucas, & Nielsen, 2017). That is, for each cohort we created two groups that were highly similar with respect to covariates, but differed on bullying perpetration. We report results for variable ratio matching because, compared to fixed ratio matching, more data could be used. As no strict cutoffs exist with regard to when sufficient balance is achieved, we examined frontier plots and computed analyses on samples in which Mahalanobis imbalance was reduced by two-third compared to the full sample. Covariates were included in the regression models on the matched samples to account for remaining bias.

Model building proceeded as follows: we first examined the prediction of health, number of partners, number of children (all cohorts), and age at first sex (TRAILS only) by bullying perpetration. Next, we computed moderation models for BCS70 and TRAILS with social status and victimization as moderators. To this end, we created interaction terms (bullying perpetration \times social status, bullying perpetration \times bullying victimization), which were entered into the prediction models alongside the main effects. We also entered the interaction term between bullying victimization and social status into the models, thereby controlling for the possibility that one interaction effect confounded the other (Keller, 2014).

Linear regression models were computed for health (all cohorts) and age at first sex (TRAILS). Poisson regressions were computed for number of partners and number of children in NCDS and BCS70. A logistic regression model was computed for number of children in TRAILS. We used robust standard errors (sandwich estimator) in the linear and Poisson regressions to allow for mild violations of the homoscedasticity assumption (linear models) and mild violation of the distribution assumption that the variance equals the mean (Poisson models) (Cameron & Trivedi, 2009).

Finally, to inspect the impact of matching on results, we also estimated conventional regression models on unmatched samples, where covariates were only entered as predictors in the model. We report unstandardized regression coefficients,

confidence intervals, and p -values as estimated in the models and note in the text which associations were stable to false discovery rate correction. For this, we used the Benjamini–Hochberg method with a critical value of $p < .05$ and correction applied across outcomes, by hypothesis and separately for boys and girls, resulting in cutoffs of $p = .0167$ for the first-ranked, $p = .033$ for the second-ranked, and $p = .05$ for the third-ranked effect for NCDS and BCS70. For TRAILS, where more outcomes were examined, the cutoffs were $p = .01$ for the first-ranked, $p = .02$ for the second-ranked, $p = .03$ for the third-ranked, $p = .04$ for the fourth-ranked, and $p = .05$ for the fifth-ranked effect.

RESULTS

Table 1 provides descriptive statistics for all three cohorts. Bullies in NCDS were less healthy but had more children than nonbullies. They had grown up in greater family instability, came from lower SES backgrounds, and scored higher on childhood psychopathology. The same pattern of differences can be observed for BCS70. There, bullies were more likely to have been victimized yet also more popular as adolescents. In TRAILS, bullies were more popular and more victimized, scored higher on childhood psychopathology and lower on SES. They reported earlier pubertal timing and age at first sex.

Bullying as Predictor of Health, Mating, and Reproductive Success

NCDS. Individuals who were rated as bullies by parents and teachers at age 16 reported poorer health at age 55 and had more children (Table 2), whereas bullying was not associated with number of partners. Results were stable across model specifications and sex. Neither for men, nor for women did statistical significance change after correction for multiple testing; bullying perpetration remained a significant predictor of health and number of children.

BCS70. Bullying perpetrators reported lower health in their early forties (Table 3). Bullying perpetration was not associated with number of partners or children, for neither men nor women. Male victims of bullying reported worse health and boys who rated themselves as popular at age 16 had more children later on, but none of the interaction effects were statistically significant. When

TABLE 2
Bullying Perpetration as Predictor of Health, Number of Partners, and Number of Children in NCDS

		Health			Number of Partners			Number of Children		
		B	95% CI (B)	p	B	95% CI (B)	p	B	95% CI (B)	p
Men	BP	-0.21	-0.35/-0.07	.003	-0.02	-0.10/0.06	.645	0.12	0.03/0.20	.006
Women	BP	-0.25	-0.42/-0.07	.006	0.05	-0.05/0.15	.342	0.14	0.05/0.24	.002

Note. Presented are unstandardized regression estimates based on matched bullying perpetrators and nonperpetrators. Matching was conducted on childhood SES, family instability, childhood psychopathology and health, BMI, and pubertal timing. All matching variables were additionally entered into the model as predictors to account for remaining bias. All outcomes were modeled separately.

correction for multiple testing was applied, the association between bullying perpetration and health remained significant in models without moderators (model 1).

TRAILS. Bullying perpetration in early adolescence was associated with age at first sex in both men and women: bullies had their first sex on average 1.2 years earlier than nonbullies (Table 4). Bullying was not related to the other outcomes; neither in men nor women. For women, the association between bullying perpetration and age at first sex remained significant after correction for multiple testing. No significant interaction effects were

found, but note that especially the number of participants with children was small.

Sensitivity Analyses 1: Conventional Regression

Matching carries the problem that achieving greater balance between cases and controls comes at the cost of removing cases that are less easily matched, for instance, because they differ substantially on the combination of covariates. To examine whether results were similar when no data were lost, we re-computed regression models without matching. Results did not change for NCDS. In BCS70, the interaction between bullying and

TABLE 3
Bullying Perpetration as Predictor of Health, Number of Partners, and Number of Children in BCS70

		Health			Number of Partners			Number of Children		
		B	95% CI (B)	p	B	95% CI (B)	p	B	95% CI (B)	p
Men										
Model 1	BP	-0.30	-0.47/-0.12	.001	0.01	-0.10/0.12	.827	0.04	-0.07/0.16	.508
Model 2	BP	-0.21	-0.41/-0.002	.048	-0.01	-0.13/0.11	.848	0.02	-0.12/0.15	.816
	BV	-0.20	-0.40/-0.003	.046	-0.04	-0.17/0.09	.568	-0.04	-0.19/0.12	.648
	POP	-0.01	-0.19/0.17	.881	0.11	-0.01/0.24	.072	0.12	0.01/0.24	.042
	BP × BV	0.06	-0.36/0.48	.781	0.09	-0.22/0.41	.567	-0.04	-0.40/0.32	.826
	BP × POP	-0.38	-0.77/0.01	.056	-0.01	-0.29/0.27	.959	0.10	-0.19/0.39	.495
	BV × POP	-0.06	-0.39/0.28	.739	0.002	-0.22/0.22	.987	-0.02	-0.26/0.22	.881
Women										
Model 1	BP	-0.30	-0.50/-0.10	.004	-0.05	-0.16/0.06	.406	0.07	-0.03/0.18	.178
Model 2	BP	-0.24	-0.45/-0.03	.028	-0.06	-0.18/0.07	.395	0.05	-0.07/0.16	.417
	BV	-0.23	-0.46/0.001	.056	0.03	-0.11/0.18	.692	0.01	-0.13/0.15	.895
	POP	0.001	-0.20/0.21	.958	0.09	-0.02/0.20	.106	0.08	-0.04/0.19	.205
	BP × BV	-0.25	-1.03/0.53	.525	-0.03	-0.47/0.40	.876	0.01	-0.33/0.36	.934
	BP × POP	-0.10	-0.60/0.42	.718	0.05	-0.28/0.38	.764	0.11	-0.16/0.38	.426
	BV × POP	-0.07	-0.57/0.42	.765	-0.13	-0.43/0.18	.413	-0.02	-0.28/0.25	.888

Note. Presented are unstandardized regression estimates based on matched bullying perpetrators and nonperpetrators. Matching was conducted on childhood SES, family instability, childhood psychopathology and health, BMI, and pubertal timing. All matching variables were additionally entered into the model as predictors to account for remaining bias. BP, bullying perpetration, BV, bullying victimization, POP, popularity. Model 1 included estimation of main effects of bullying perpetration only; Model 2 included all interaction effects.

TABLE 4
Bullying Perpetration as Predictor of Health, Age at First Sex, Number of Partners, and Number of Children in TRAILS

	General Health			Physical Complaints			Age at first Sex			Number of Sexual Partners			Number of Children			
	B	95% CI (B)	p	B	95% CI (B)	p	B	95% CI (B)	p	B	95% CI (B)	p	B	95% CI (B)	p	
Men																
Model 1	BP	-0.16	-0.44/0.12	.255	-0.01	-11/0.08	.767	-1.19	-2.24/-0.14	.027	0.06	-0.16/0.27	.612	0.25	-0.98/1.47	.690
Model 2	BP	-0.01	-0.45/0.44	.972	-0.01	-14/0.12	.906	-1.54	-3.09/0.01	.052	0.04	-0.31/0.40	.816	1.28	-0.55/3.11	.170
	BV	0.18	-0.33/0.69	.487	0.08	-08/0.23	.330	-1.00	-3.08/1.09	.347	0.15	-0.21/0.51	.406	-0.82	-500.63/498.98	.997
	POP	0.07	-0.43/0.58	.776	0.06	-11/0.23	.488	-2.50	-4.35/-0.64	.009	0.38	-0.001/0.77	.051	0.55	-1.46/2.55	.592
	BP × BV	-0.47	-1.26/0.32	.240	-0.03	-25/0.18	.778	1.06	-1.58/3.70	.431	0.07	-0.44/0.58	.789	-1.25	-608.95/606.45	.997
	BP × POP	0.19	-0.59/0.98	.624	-0.01	-30/0.29	.961	2.11	-0.52/4.73	.114	-0.38	-1.02/0.25	.237	-2.21	-5.53/1.10	.189
	BV × POP	-0.46	-1.63/0.70	.430	-0.16	-48/0.16	.327	-0.47	-3.66/2.71	.769	-0.07	-0.88/0.75	.866	1.54	-784.89/787.97	.997
Women																
Model 1	BP	0.05	-0.34/0.44	.806	0.04	-0.08/0.17	.515	-1.31	-2.27/-0.36	.007	-0.004	-0.22/0.22	.973	0.25	-0.98/1.47	.690
Model 2	BP	0.32	-0.16/0.80	.189	0.06	-0.11/0.23	.480	-2.05	-3.39/-0.71	.003	-0.16	-0.53/0.22	.415	1.28	-0.55/3.11	.170
	BV	0.24	-0.45/0.92	.499	-0.04	-0.28/0.19	.708	-0.31	-2.32/1.69	.758	0.02	-0.35/0.38	.933	-0.82	-500.63/498.98	.997
	POP	0.21	-0.40/0.81	.504	-0.05	-0.23/0.14	.607	-1.20	-2.63/0.23	.998	-0.06	-0.44/0.32	.744	0.55	-1.46/2.55	.592
	BP × BV	-0.41	-1.21/0.38	.310	-0.01	-0.32/0.30	.939	1.66	-0.88/4.19	.200	-0.06	-0.58/0.47	.826	-1.25	-608.95/606.45	.997
	BP × POP	-0.52	-1.39/0.35	.238	-0.01	-0.30/0.28	.948	1.17	-0.76/3.09	.234	0.35	-0.19/0.89	.201	-2.21	-5.53/1.10	.189
	BV × POP	0.18	-0.91/1.26	.750	0.10	-0.29/0.48	.627	0.87	-2.00/3.74	.551	0.06	-0.53/0.65	.846	1.54	-784.89/787.97	.997

Note. Presented are unstandardized regression estimates based on matched bullying perpetrators and nonperpetrators. Matching was conducted on childhood SES, family instability, childhood psychopathology and health, BMI, and pubertal timing. All matching variables were additionally entered into the model as predictors to account for remaining bias. BP, bullying perpetration; BV, bullying victimization; POP, popularity. Model 1 included estimation of main effects of bullying perpetration only; Model 2 included all interaction effects. We do not include results pertaining to numbers of children in men as confidence intervals were extremely large, pointing at unreliable estimates.

popularity predicted health in male participants ($B = -0.36$, 95% CI $[-0.66, -0.06]$, $p = .019$; though note the cutoff for multiple testing), yet opposite to expectations: popular bullies fared *worse* with respect to later health. Not of primary interest to the present study yet different from the matched analyses, popular boys in BCS70 reported more partners ($B = 0.11$, 95% CI $[0.04, 0.19]$, $p = .002$). Popular girls also had more partners ($B = 0.08$, 95% CI $[0.02, 0.14]$, $p = .006$) and more children ($B = 0.07$, 95% CI $[0.01, 0.13]$, $p = .017$; though note the cutoff for multiple testing). In TRAILS, popular boys reported more sexual partners ($B = 0.33$, 95% CI $[0.15, 0.50]$, $p < .001$).

Sensitivity Analyses 2: Including One Interaction Term at a Time

To explore whether overfitting masked interaction effects in models with multiple interaction terms, we re-estimated models with interaction effects between bullying and popularity and bullying and victimization entered separately. In BCS70, the interaction between bullying and popularity in predicting health among men was statistically significant (before multiple testing correction) when matched data were used ($B = -0.39$, 95% CI $[-0.74, -0.03]$, $p = .033$) and also when non-matched data were used ($B = -0.38$, 95% CI $[-0.65, -0.12]$, $p = .005$), again suggesting that popular bullies fared *worse* with respect to later health than nonbullies and unpopular bullies. No other interactions reached statistical significance in these models.

DISCUSSION

Bullying has played an important role in the developmental literature and in educational policy in the last few years. This surge in attention was fueled by a plethora of studies on maladjustment of victims. Negative outcomes have also been ascribed to perpetrators, but the scientific evidence here is thinner. In fact, some longitudinal studies leave doubt as to whether bullying others actually carries risk for maladjustment (Copeland et al., 2013; Wolke et al., 2013). Bullying research inspired by evolutionary theory has even suggested that bullies might reap benefits in the form of better health, access to partners, and reproductive opportunities. Here, we tested whether bullying was indeed linked to these outcomes in three cohorts.

Across the older cohorts (NCDS and BCS70), bullies showed *worse* health outcomes in middle

adulthood. At first sight, this finding is *not* supportive of the hypothesis that bullying perpetrators are exposed to less stress given their rank in the social pecking order and thus should theoretically reap the benefits in form of better health (Volk et al., 2012). It appears that decreased stress among bullies is linked to better health in the short term—as supported by a study into differences by bullying status in increases in systemic inflammation levels from adolescence to early adulthood (Copeland et al., 2014)—but that this beneficial effect does not last for decades. In fact, bullying perpetrators engage in health-adverse behaviors (Ttofi et al., 2016) and experience more stress—though not necessarily worse health—as adults (Matthews et al., 2017). It is possible that negative health outcomes emerge later in life, which might explain why we did not observe this association in TRAILS. For instance, bullying perpetrators might continue to aggress against work colleagues and partners (Farrell & Vaillancourt, 2019; Matthiesen & Einarsen, 2007), thus jeopardize employment, friendships, and relationships and end up unhappy and unhealthy as adults. Bullying perpetration has also been associated with substance use later on (Vrijen et al., 2021), which has negative effects on physical health as well. Not detecting this association in TRAILS might mean that bullying perpetration in this cohort is still ongoing and health benefits resulting from lower social stress are continued to be reaped for the time being.

Notably, whereas health was worse among bullying perpetrators in NCDS and BCS70, NCDS participants had more offspring, the clearest indicator of evolutionary benefit. An association with a greater number of offspring was also observed among BCS70 when bullies and nonbullies were compared prior to matching. Bullies in the TRAILS sample engaged in sexual behavior earlier than nonbullies. Whereas early sex tends to be seen as risk behavior in developmental terms, it widens the span of years for reproduction and is thus considered beneficial from an evolutionary perspective. In this regard, the TRAILS results match those for NCDS and the unadjusted ones for BCS70.

It is worth keeping in mind that it might not be bullying in and of itself that conveys a reproductive advantage but that unexamined third variables explain both. Life-history theory, for instance, would suggest that traits and behaviors that represent fast life strategy and allocation of resources to reproduction increase someone's likelihood to bully others as well as lower age at first sex and more offspring. A life-history perspective would also

explain long-term links between bullying and the other outcomes examined in this study as individuals following a fast life-history strategy allocate resources to reproduction rather than somatic fitness. Bullies might forego good health in order to reap reproductive benefits, which is ultimately adaptive because transmission of genes is more important in an evolutionary sense than being healthy in old age. As such, negative health outcomes for bullying perpetrators would not necessarily mean that bullying is not at all adaptive because reproducing early and manifold is more important than preserving good health. A reproductive-versus-somatic trade-off would explain why bullying perpetrators enjoy greater reproductive success but suffer from worse health later on. Unfortunately, NCDS is the only cohort in our study where reproductive activity can be assumed to be more or less concluded, certainly among the female participants. If indeed the number of offspring is also higher among bullying perpetrators in BCS70 once this cohort has completed reproduction, that is, if NCDS results are replicated in another cohort, it would be interesting to study whether bullies who had more offspring are worse off in terms of health; thus whether short-term benefits in favor of long-term costs can be observed.

Perspectives on bullying that are not informed by evolutionary theory would suggest that bullying perpetration and earlier sex—for which we found a link in TRAILS—are both expressions of externalizing maladjustment, a view that in our data is supported by the higher levels of child psychopathology among bullies than nonbullies. Age at first sex and the bullying perpetration assessment will have been temporally closer to one another than bullying and some of the other outcomes in TRAILS, which underlines the possibility that both are concurrent expressions of an underlying construct such as externalizing behavior. It will be important to follow up on the BCS70 and TRAILS participants at an age similar to that of NCDS participants now to see whether divergent results across the samples used here indicate unstable, nonreplicable associations or have substantive meaning.

One might doubt whether are earlier age at first sex and number of partners and children are meaningful indicators of success and advantage in contemporary Western societies where slow life strategies dominate (Twenge & Park, 2019). At least within the developmental literature on adolescence and young adulthood, having many partners is considered an expression of maladjustment and

risky sexual behavior and promiscuity in adulthood is often seen as a symptom of psychopathology. Contextual conditions might also be implicated in this link, such as norms in the peer group that favor both bullying and risky sexual behavior. Though this perspective suggests that it is not meaningful to study number of partners and children as solely positive outcomes, they still indicate reproductive success in an evolutionary sense. In other words, what is adaptive in an ultimate sense (maximized reproduction) might not be desirable in a proximate sense, given the stigma attached to families with greater number of offspring and societal norms surrounding numbers of sexual partners.

Limitations

Perhaps the most substantial limitation of this study is the bullying perpetration measurement. Though it is remarkable that researchers assessed bullying when child psychology had hardly picked up on the topic and its negative developmental outcomes, single-item assessments are clearly not optimal. Parents and teachers have a limited view on who engages in bullying (Ahn, Rodkin, & Gest, 2013) – and will have had so even more at a time when bullying did not receive the attention from educational policymakers and media as it does today. As a consequence, measures of bullying in NCDS and BCS70 might lack validity and will also not have included nonovert forms of bullying. Moreover, bullying by popular, high-status adolescents might not have been viewed by adults as the destructive behavior it is for victims and noninvolved peers and parents might generally be unaware of the behavior of their children at school unless informed and might have responded with other target of bullying in mind. Of note, however, the proportion of bullying perpetrators in NCDS and BCS70 is similar to that found in studies that used multi-item self-report measures (Espelage, Van Ryzin, & Holt, 2018). Moreover, whereas more recent studies provide participants with a definition of bullying that ideally encompasses its different dimensions (Kaufman, Huitsing, & Veenstra, 2020) such detail was not provided to the three samples on which our analyses are based. It might thus be that the outcomes only hold for perpetrators who engage in visible, direct forms of bullying behavior, such as physical and verbal bullying. Next to using bullying information from different reporters in TRAILS versus NCDS and BCS70, assessments were also done at a slightly younger

age in TRAILS. Bullying is related to status more so in adolescence than in childhood which might have meant that associations with outcomes would be more pronounced in the cohorts where the assessment was conducted in mid- rather than early adolescence.

It could also be argued that what we conceptualized as bullying is actually general aggression, or understood by parents and teachers in NCDS and BCS70 as tapping into aggression more generally. Both bullying and aggression are used strategically to obtain social dominance to control resources (Hawley, 1999) and to increase reproductive success (Lindenfors & Tullberg, 2011; Vaillancourt, 2005). As such, the evolutionary propositions would to some extent be similar but lower health in adulthood would be expected for aggressive behavior and not necessarily, from an evolutionary perspective, for bullying perpetration, even though bad health might be acceptable in light of successful reproduction. We corrected for childhood psychopathology to account for this at least partly.

Next to the use of a single item to measure bullying, the present study should be interpreted in light of other methodological challenges: First, attrition was selective and, across samples, bullying perpetration predicted missing data on outcomes. Whereas multiple imputation provides a more advanced way to deal with missing data than case deletion or mean imputation, it would be important to see whether the results hold on more complete data.

Second, we dichotomized ordinal bullying perpetration scores for NCDS and BCS and proportion scores for TRAILS. This was done to harmonize bullying perpetration measures as much as possible across samples, avoid skewed distributions, and to enable the matching procedure but also meant a loss of information. We chose cutoff scores for dichotomization in a way that would return reasonably comparable group sizes and applied this strategy also to victimization and popularity. It is not optimal that these cutoffs are therefore based on methodological rather than substantive considerations. One might ask whether results would have looked different if only those individuals with more extreme scores would have been included in the bully matching group. It is feasible that associations would be more pronounced (even worse health outcomes, more partners, more children) but this would have led to a vastly smaller group than its counterpart (nonbully), which would have increased the risk for model instability.

Third, the matching procedure allowed for a range of relatively subjective decisions, for instance between fixed and variable ratio matching and with respect to pruning (e.g., more pruning leads to better matching but to smaller samples). As developmental researchers increasingly use advanced methods, close collaboration between developers of such analyses packages and applied researchers is needed to ensure user friendliness and correct application.

Fourth, three samples from different periods were included which means that differences in effects might be explained by cohort effects. NCDS represents the first generation to grow up with the birth control pill being widely available but maybe not yet as accepted across all groups in the UK population. Tentatively, this might mean that number of children was less controlled by women in this cohort than in BCS70 and TRAILS. Similarly, covariates such as SES might have exerted a different influence on outcomes on those from the earlier compared to later-assessed cohorts and childhood psychopathology might have carried a greater stigma in the older cohorts.

Fifth, differences in measurement across the cohorts, and especially for TRAILS in comparison to NCDS and BCS70, might explain discrepancies in results. An example for this is the health assessment, which was kept broad in NCDS and BCS70 but referred explicitly to physical health in TRAILS. Another example concerns the number of partners, which reflects number of partners with whom the participant has cohabited for at least 1 month in NCDS and BCS70 whereas this variable refers to number of sexual partners in TRAILS. The former represents an imperfect assessment from an evolutionary standpoint, where number of sexual partners would have been the preferred measure. The value of the British cohorts in terms of length and inclusion of a bullying assessment long before the topic entered the mind of researchers, however, outweigh this limitation in our view.

Sixth, our central aim was to test Volk's hypothesis of bullying as evolutionary adaptation (Volk et al., 2012) and we selected three cohorts where outcomes were assessed at different ages. Whereas we examined moderation by popularity and victimization – assessed at the same time as bullying perpetration, we did not explore potential mechanisms that might carry the effects of adolescent bullying into (late) adulthood. Our reasoning implicitly suggested that bullies retain a high status, but this remains to be tested. Another possibility might be that bullies who are successful in

dating in adolescence and early adulthood build up and retain self-confidence when it comes to establishing sexual relationships and are able to build larger social networks which again adds opportunities to meet a potential (sexual) partner. Adolescents who frequently change their dating partners might settle down later and thus accumulate a greater number of partners even into adulthood. At present, the lack of longitudinal studies on stability of bullying status hinders any rigorous testing of such mechanisms. In a similar vein, it is difficult to derive immediate practical implications from the findings presented here. Naturally, bullying prevention needs to remain on the agenda even if only for victims, who carry the greatest plight. Reducing bullying in schools, however, will hardly eradicate evolutionary advantages for high-status individuals and high status among adolescents is linked to aggressive behavior. From a life-history strategy perspective, one might want to attempt to reduce antecedents for fast life strategies, such as growing up in harsh and unpredictable environments (Belsky, Schlomer, & Ellis, 2012), but it is somewhat doubtful that these are tasks for antibullying pre- and intervention programs. In that sense, the results of this study should be seen as furthering our understanding of long-term correlates of bullying but cannot immediately be translated into policy or practical implications.

Finally, we preregistered analyses but realized that not all premade analytic decisions were optimally suited. For instance, we intended to follow previous conceptualizations of bullying perpetration in TRAILS (Veenstra et al., 2005) but discovered that this would lead to too few bullying perpetrator cases to meaningfully use in matching. Moreover, we had not anticipated that so few children had been born to TRAILS participants yet, which made analyses with number of children as a count variable unreliable. These deviations from the preregistration highlight that some decisions are easier when the data are known.

CONCLUSION

Evolutionary developmental theory suggests an advantage for bullies compared to nonbullies, but our findings from three longitudinal cohorts show that bullies were less healthy in middle adulthood. Bullies had more children when assessed later in adulthood, suggesting greater reproductive success and especially female bullying perpetrators engaged in sexual activity at an earlier age which is considered reproductively beneficial but

developmentally a risk. From a life-history perspective, this pattern of findings might be indicative of a somatic-versus-reproductive trade-off whereby bullies follow a fast life strategy that implies investing in reproductive opportunities even if those come at health costs later on. As such, our data lend partial support to an evolutionary advantage of bullying perpetration but understanding concrete mechanisms requires further research with robust assessments of bullying, ideally from multiple reporters and using refined measures.

REFERENCES

- Achenbach, T. M. & Rescorla, L. A. (2001). Manual for the ASEBA school-age forms & profiles: Child behavior checklist for ages 6-18, teacher's report form, youth self-report: An integrated system of multi-informant assessment. University of Vermont, Research Center for Children Youth & Families.
- Ahn, H.-J., Rodkin, P. C., & Gest, S. (2013). Teacher-student agreement on "bullies and kids they pick on" in elementary school classrooms: Gender and grade differences. *Theory into Practice*, 52(4), 257-263. <https://doi.org/10.1080/00405841.2013.829728>
- Archer, J., & Benson, D. (2008). Physical aggression as a function of perceived fighting ability and provocation: An experimental investigation. *Aggressive Behavior*, 34(1), 9-24. <https://doi.org/10.1002/ab.20179>
- Barnard, J., & Rubin, D. B. (1999). Small-sample degrees of freedom with multiple imputation. *Biometrika*, 86(4), 948-955. <https://doi.org/10.1093/biomet/86.4.948>
- Belsky, J., Schlomer, G. L., & Ellis, B. J. (2012). Beyond cumulative risk: Distinguishing harshness and unpredictability as determinants of parenting and early life history strategy. *Developmental Psychology*, 48(3), 662-673.
- Cameron, A. C., & Trivedi, P. K. (2009). *Microeconometrics using Stata*. Stata Press.
- Caspi, A., Moffitt, T. E., Thornton, A., Freedman, D., Amell, J. W., Harrington, H., ... Silva, P. A. (1996). The life history calendar: A research and clinical assessment method for collecting retrospective event-history data. *International Journal of Methods in Psychiatric Research*, 6, 101-114. [https://doi.org/10.1002/\(SICI\)1234-988X\(199607\)6:2<101::AID-MPR156>3.3.CO;2-E](https://doi.org/10.1002/(SICI)1234-988X(199607)6:2<101::AID-MPR156>3.3.CO;2-E)
- Connolly, J., Pepler, D., Craig, W., & Taradash, A. (2000). Dating experiences of bullies in early adolescence. *Child Maltreatment*, 5(4), 299-310. <https://doi.org/10.1177/1077559500005004002>
- Copeland, W. E., Wolke, D., Angold, A., & Costello, E. J. (2013). Adult psychiatric outcomes of bullying and being bullied by peers in childhood and adolescence. *JAMA Psychiatry*, 70(4), 419-426. <https://doi.org/10.1001/jamapsychiatry.2013.504>
- Copeland, W. E., Wolke, D., Lereya, S. T., Shanahan, L., Worthman, C., & Costello, E. J. (2014). Childhood

- bullying involvement predicts low-grade systemic inflammation into adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 7570–7575. <https://doi.org/10.1073/pnas.1323641111>
- Dane, A. V., Marini, Z. A., Volk, A. A., & Vaillancourt, T. (2017). Physical and relational bullying and victimization: Differential relations with adolescent dating and sexual behavior. *Aggressive Behavior*, 43(2), 111–122. <https://doi.org/10.1002/ab.21667>
- Dijkstra, J. K., Cillessen, A. H., Lindenberg, S., & Veenstra, R. (2009). Same-gender and cross-gender likeability: Associations with popularity and status enhancement: The TRAILS study. *Journal of Early Adolescence*, 30(6), 773–802. <https://doi.org/10.1177/0272431609350926>
- Elliott, J., & Shepherd, P. (2006). Cohort Profile: 1970 British Birth Cohort (BCS70). *International Journal of Epidemiology*, 35(4), 836–843. <https://doi.org/10.1093/ije/dyl174>
- Ely, M., Richards, M. P. M., Wadsworth, M. E. J., & Elliott, B. J. (1999). Secular changes in the association of parental divorce and children's educational attainment – Evidence from three British birth cohorts. *Journal of Social Policy*, 28(3), 437–455. <https://doi.org/10.1017/S0047279499005693>
- Espelage, D. L., Van Ryzin, M. J., & Holt, M. K. (2018). Trajectories of bully perpetration across early adolescence: Static risk factors, dynamic covariates, and longitudinal outcomes. *Psychology of Violence*, 8(2), 141–150. <https://doi.org/10.1037/vio0000095>
- Farrell, A. H., & Vaillancourt, T. (2019). Temperament, bullying, and dating aggression: Longitudinal associations for adolescents in a romantic relationship. *Evolutionary Psychology*, 17(2). <https://doi.org/10.1177/1474704919847450>
- Ganzeboom, H. B., & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations. *Social Science Research*, 25(3), 201–239. <https://doi.org/10.1006/ssre.1996.0010>
- Garza, J. R., Glenn, B. A., Mistry, R. S., Ponce, N. A., & Zimmerman, F. J. (2017). Subjective social status and self-reported health among US-born and immigrant Latinos. *Journal of Immigrant and Minority Health*, 19(1), 108–119. <https://doi.org/10.1007/s10903-016-0346-x>
- Hancock, M. (2016). British Cohort Study (1970) Partnership Histories (1986–2013). Centre for Longitudinal Studies: Institute of Education, University of London.
- Hancock, M. (2017). National Child Development Study Partnership Histories (1974–2013) (p. 21). Centre for Longitudinal Studies, Institute of Education, London.
- Hawley, P. H. (1999). The ontogenesis of social dominance: A strategy-based evolutionary perspective. *Developmental Review*, 19(1), 97–132. <https://doi.org/10.1006/drev.1998.0470>
- Kaufman, T. M., Huitsing, G., & Veenstra, R. (2020). Refining victims' self-reports on bullying: Assessing frequency, intensity, power imbalance, and goal-directedness. *Social Development*, 29(2), 375–390. <https://doi.org/10.1111/sode.12441>
- Keller, M. C. (2014). Gene x environment interaction studies have not properly controlled for potential confounders: The problem and the (simple) solution. *Biological Psychiatry*, 75(1), 18–24. <https://doi.org/10.1016/j.biopsych.2013.09.006>
- King, G., Lucas, C., & Nielsen, R. A. (2017). The balance-sample size frontier in matching methods for causal inference. *American Journal of Political Science*, 61(2), 473–489. <https://doi.org/10.1111/ajps.12272>
- Koh, J.-B., & Wong, J. S. (2017). Survival of the fittest and the sexiest: Evolutionary origins of adolescent bullying. *Journal of Interpersonal Violence*, 32(17), 2668–2690. <https://doi.org/10.1177/0886260515593546>
- Kretschmer, T., Veenstra, R., Branje, S., Reijneveld, S. A., Meeus, W. H. J., Deković, M., ... Oldehinkel, A. J. (2018). How competent are adolescent bullying perpetrators and victims in mastering normative developmental tasks in early adulthood? *Journal of Abnormal Child Psychology*, 46(1), 41–56. <https://doi.org/10.1007/s10802-017-0316-3>
- Lindfors, P., & Tullberg, B. S. (2011). Evolutionary aspects of aggression: The importance of sexual selection. *Advances in genetics*, 75, 7–22. <https://doi.org/10.1016/B978-0-12-380858-5.00009-5>
- Matthews, K. A., Jennings, J. R., Lee, L., & Pardini, D. A. (2017). Bullying and being bullied in childhood are associated with different psychosocial risk factors for poor physical health in men. *Psychological Science*, 28(6), 808–821. <https://doi.org/10.1177/0956797617697700>
- Matthiesen, S. B., & Einarsen, S. (2007). Perpetrators and targets of bullying at work: Role stress and individual differences. *Violence and Victims*, 22(6), 735–753. <https://doi.org/10.1891/088667007782793174>
- Mundt, M. P., & Zakletskaia, L. I. (2014). That's what friends are for: Adolescent peer social status, health-related quality of life and healthcare costs. *Applied Health Economics and Health Policy*, 12(2), 191–201. <https://doi.org/10.1007/s40258-014-0084-y>
- Oldehinkel, A. J., Rosmalen, J. G., Buitelaar, J. K., Hoek, H. W., Ormel, J., Raven, D., ... Hartman, C. A. (2015). Cohort Profile Update: The TRacking Adolescents' Individual Lives Survey (TRAILS). *International Journal of Epidemiology*, 44(1), 76–76n. <https://doi.org/10.1093/ije/dyu225>
- Power, C., & Elliott, J. (2006). Cohort profile: 1958 British birth cohort (National Child Development Study). *International Journal of Epidemiology*, 35(1), 34–41. <https://doi.org/10.1093/ije/dyi183>
- Provenzano, D. A., Dane, A. V., Farrell, A. H., Marini, Z. A., & Volk, A. A. (2018). Do bullies have more sex? The role of personality. *Evolutionary Psychological Science*, 4(3), 221–232. <https://doi.org/10.1007/s40806-017-0126-4>
- Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health and behaviour*. Longman Publishing Group.
- Savickaitė, R., Dijkstra, J. K., Kreager, D., Ivanova, K., & Veenstra, R. (2020). Friendships, perceived popularity,

- and adolescent romantic relationship debut. *The Journal of Early Adolescence*, 40(3), 377–399. <https://doi.org/10.1177/0272431619847530>
- Schoeler, T., Duncan, L., Cecil, C. M., Ploubidis, G. B., & Pingault, J.-B. (2018). Quasi-experimental evidence on short- and long-term consequences of bullying victimization: A meta-analysis. *Psychological Bulletin*, 144, 1229–1246. <http://dx.doi.org/10.1037/bul0000171>
- Sentse, M., Kretschmer, T., & Salmivalli, C. (2015). The longitudinal interplay between bullying, victimization, and social status: Age-related and gender differences. *Social Development*, 24(3), 659–677. <https://doi.org/10.1111/sode.12115>
- Sweeting, H., & Hunt, K. (2014). Adolescent socioeconomic and school-based social status, health and well-being. *Social Science & Medicine*, 121, 39–47. <https://doi.org/10.1016/j.socscimed.2014.09.037>
- Takizawa, R., Maughan, B., & Arseneault, L. (2014). Adult health outcomes of childhood bullying victimization: Evidence from a five-decade longitudinal British birth cohort. *American Journal of Psychiatry*, 171(7), 777–784. <https://doi.org/10.1176/appi.ajp.2014.13101401>
- Ttofi, M. M., Farrington, D. P., Lösel, F., Crago, R. V., & Theodorakis, N. (2016). School bullying and drug use later in life: A meta-analytic investigation. *School Psychology Quarterly*, 31(1), 8. <https://doi.org/10.1037/spq0000120>
- Twenge, J. M., & Park, H. (2019). The decline in adult activities among US adolescents, 1976–2016. *Child Development*, 90(2), 638–654. <https://doi.org/10.1111/cdev.12930>
- Vaillancourt, T. (2005). In R. E. Tremblay, W. W. Hartup, & J. Archer (Eds.), *Indirect aggression among humans: Social construct or evolutionary adaptation?* Developmental Origins of Aggression. (pp. 158–177). New York: Guilford.
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1–67. <https://doi.org/10.18637/jss.v045.i03>
- van der Ploeg, R., Steglich, C., & Veenstra, R. (2020). The way bullying works: How new ties facilitate the mutual reinforcement of status and bullying in elementary schools. *Social Networks*, 60, 71–82. <https://doi.org/10.1016/j.socnet.2018.12.006>
- Veenstra, R., Lindenberg, S., Oldehinkel, A. J., De Winter, A. F., & Ormel, J. (2006). Temperament, environment, and antisocial behavior in a population sample of preadolescent boys and girls. *International Journal of Behavioral Development*, 30(5), 422–432. <https://doi.org/10.1177/0165025406071490>
- Veenstra, R., Lindenberg, S., Oldehinkel, A. J., Winter, A. F. D., Verhulst, F. C., & Ormel, J. (2005). Bullying and victimization in elementary schools: A comparison of bullies, victims, bully/victims, and uninvolved preadolescents. *Developmental Psychology*, 41(4), 672–682. <https://doi.org/10.1037/0012-1649.41.4.672>
- Volk, A. A., Camilleri, J. A., Dane, A. V., & Marini, Z. A. (2012). Is adolescent bullying an evolutionary adaptation? *Aggressive Behavior*, 38(3), 222–238. <https://doi.org/10.1002/ab.21418>
- Volk, A. A., Dane, A. V., Marini, Z. A., & Vaillancourt, T. (2015). Adolescent bullying, dating, and mating testing an evolutionary hypothesis. *Evolutionary Psychology*, 13(4), 147470491561390. <https://doi.org/10.1177/1474704915613909>
- Volk, A. A., Della Cioppa, V., Earle, M., & Farrell, A. H. (2015). Social competition and bullying: An adaptive socioecological perspective. In V. Zeigler-Hill, L. L. M. Welling, & T. K. Shackelford (Eds.), *Evolutionary perspectives on social psychology* (pp. 387–399). Switzerland: Springer International Publishing.
- Von Hippel, P. T. (2009). How to impute interactions, squares, and other transformed variables. *Sociological Methodology*, 39(1), 265–291. <https://doi.org/10.1111/j.1467-9531.2009.01215.x>
- Vrijen, C., Wiertsema, M., Ackermans, M. A., van der Ploeg, R., & Kretschmer, T. (2021). Childhood and adolescent bullying perpetration and later substance use: A meta-analysis. *Pediatrics*, 147(3), e2020034751. <https://doi.org/10.1542/peds.2020-034751>
- White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, 30(4), 377–399. <https://doi.org/10.1002/sim.4067>
- Wolke, D., Copeland, W. E., Angold, A., & Costello, E. J. (2013). Impact of bullying in childhood on adult health, wealth, crime, and social outcomes. *Psychological Science*, 24(10), 1958–1970. <https://doi.org/10.1177/0956797613481608>
- Wood, N., Bann, D., Hardy, R., Gale, C., Goodman, A., Crawford, C., & Stafford, M. (2017). Childhood socioeconomic position and adult mental wellbeing: Evidence from four British birth cohort studies. *PLoS One*, 12(10), e0185798. <https://doi.org/10.1371/journal.pone.0185798>
- Zwaan, M., Dijkstra, J. K., & Veenstra, R. (2013). Status hierarchy, attractiveness hierarchy and sex ratio: Three contextual factors explaining the status–aggression link among adolescents. *International Journal of Behavioral Development*, 37(3), 211–221. <https://doi.org/10.1177/0165025412471018>

Supporting Information

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

Table S1. Attrition analyses for NCDS.

Table S2. Attrition analyses for BCS70.

Table S3. Attrition analyses for TRAILS.