Early Adolescent Health Problems, School Performance, and Upper Secondary Educational Pathways : A Counterfactual-Based Mediation Analysis

Mikkonen, Janne

2021-03

Mikkonen , J , Remes , H , Moustgaard , H & Martikainen , P 2021 , ' Early Adolescent Health Problems, School Performance, and Upper Secondary Educational Pathways : A Counterfactual-Based Mediation Analysis ' , Social Forces , vol. 99 , no. 3 , pp. 1146-1175 . https://doi.org/10.1093/s

http://hdl.handle.net/10138/344158
https://doi.org/10.1093/sf/soaa048

acceptedVersion

Early-adolescent health problems, school performance, and upper-secondary educational pathways – a counterfactual-based mediation analysis

Janne Mikkonen[1], Hanna Remes[1], Heta Moustgaard[1] & Pekka Martikainen[1,2,3]

[1] Population Research Unit, Faculty of Social Sciences, University of Helsinki, Finland
[2] Center for Health Equity Studies (CHESS), Department of Public Health Sciences, Stockholm University, Stockholm, Sweden
[3] Max Planck Institute for Demographic Research, Rostock, Germany

Corresponding author:

Janne Mikkonen
Address: Population Research Unit, Faculty of Social Sciences, University of Helsinki, PO Box 18 (Unioninkatu 35), FIN-00014, Finland
Email: janne.mk.mikkonen@helsinki.fi

Keywords:

Education, Health problems, Non-completion, Track choice, Adolescence, Mediation analysis

Acknowledgements:

This study was supported by the Academy of Finland (grant number: 294861) and the Strategic Research Council of Finland (grant number: 303615). The authors wish to thank Statistics Finland and the National Institute for Health and Welfare for granting access to the data.

Early-adolescent health problems, school performance, and upper-secondary educational pathways – a counterfactual-based mediation analysis

Abstract

Early health problems predict lower educational attainment, but it remains unclear whether this is due to health problems weakening school performance or due to other mechanisms operating above and beyond school performance. We employed counterfactual-based mediation analysis on a register-based sample of Finnish adolescents born in 1988-1993 (n=73,072) to longitudinally assess the direct (unexplained by school performance, as measured by grade point average) and indirect (pure mediation and mediated interaction via school performance) effects of early-adolescent somatic and mental health problems on the non-completion of upper-secondary education and track choice (vocational vs. general). Mental disorders were associated with the largest increases in both non-completion and choosing the vocational track, but somatic conditions also showed small but robust associations. Weakened school performance mediated up to one third of the differences in non-completion and around half of the differences in track choice. When the same analyses were conducted within sibships, the total effects of health problems on educational pathways were weaker, but the contribution of school performance remained similar. In counterfactual simulations that assigned everyone an above-median school performance—i.e., eradicating below-median school performance—about 20-40 percent of the effects of mental disorders on educational pathways remained. Our results suggest that while impaired school performance is an important component in health-related selection to education, it does not fully explain the shorter and less academically oriented educational careers of adolescents with health problems. These adolescents may benefit from additional educational support regardless of their formal school performance.

1 Introduction

The widely demonstrated salience of education for both individual life chances and society at large makes a case for tackling obstacles in educational progression (McMahon and Oketch 2013). A report by the World Health Organization summarized that poor childhood health is extensively associated with worse school performance and lower educational attainment in developed countries (Suhrcke and de Paz Nieves 2011), and similar conclusions can be drawn from other systematic reviews (Hale, Bevilacqua, and Viner 2015; Melkevik et al. 2016; Esch et al. 2014). Despite the relatively large amount of correlational evidence on health–schooling relationships, the general view of the processes at work has remained blurred due to the considerable data requirements involved in measuring both health and educational outcomes with an adequately long follow-up.

A plausible—but so far rarely examined—factor contributing to the lower educational attainment of those with early health problems is weakened school performance. Both physical and mental health problems have predicted weaker performance in previous studies (Crump et al. 2013; McLeod, Uemura, and Rohrman 2012; Eide, Showalter, and Goldhaber 2010), and prior grades and test scores are a common component of admission criteria in upper-secondary and tertiary education worldwide (OECD 2004, 17–20; 2018b, 2–3). Moreover, a more nuanced view of health-related selection to education could be achieved by extending focus from vertical stratification in educational outcomes, such as years of education or school completion and dropout (Mikkonen et al. 2018; Sagatun et al. 2014; Haas and Fosse 2008), to horizontal stratification, i.e., track choice.

This study uses Finnish longitudinal register data that includes precise measurements of healthcare use, school performance at the end of compulsory school, and upper-secondary degrees for a nationally representative sample of 73,072 children to assess the role of school performance in health-related selection to education. We examine whether adolescents experiencing somatic conditions or mental disorders at ages 10–16 are more likely to not complete upper-secondary education by age 23, and whether those completing an upper-secondary degree are more likely to

choose the vocational track that prepares people to work in manual or lower non-manual jobs and is academically less demanding than its alternative, the general track. The principal contribution of our study is to employ g-computation for counterfactual-based mediation analysis to decompose the total effects of health problems on upper-secondary education into three components: *pure direct effect* unexplained by weakened school performance, *pure indirect effect* via weakened school performance, and *mediated interaction effect* that identifies the possible differential effects of poor school performance between adolescents with and without health problems. We estimate these effects both in models adjusting for a large number of confounders and in sibling fixed-effects models that adjust for all unobserved confounding shared within sibships. As a policy-relevant illustration of the part unexplained by school performance, we further calculated *controlled direct effects* that simulate a hypothetical intervention of providing everyone with above-median school performance.

1.1 Background

Stunting of educational careers due to health problems is receiving increasing theoretical and empirical attention (Layte and McCrory 2013). Most evidence on the association between health problems and educational outcomes among young people relies on measures of self-rated health or self-reported symptoms and health behaviors (Margot Jackson 2009; Huurre et al. 2006; Lê, Diez Roux, and Morgenstern 2013; Brekke 2015). The smaller cluster of studies measuring the effects of specific health conditions or larger groups of serious health conditions have typically highlighted the particular significance of child and adolescent mental disorders as predictors of weaker school performance, school dropout, or lower educational attainment in general (Breslau et al. 2008; Mikkonen et al. 2018; Isohanni et al. 2001; Kessler et al. 1995; McLeod and Fettes 2007; McLeod and Kaiser 2004). These associations have been the strongest in the case of rare, psychotic disorders as well as conduct and attention-deficit disorders (Johnson et al. 1999; Fried et al. 2016; Evensen et al. 2016; Isohanni et al. 2001; Goulding, Chien, and Compton 2010; Frissen et al. 2015),

whereas the impact of anxiety and depression on educational outcomes, independent of co-occurring externalizing disorders, is more debated (Miech et al. 2002; Needham 2009; Melkevik et al. 2016; Evensen et al. 2016; Fergusson and Woodward 2002; McLeod and Kaiser 2004; Fletcher 2010).

As for somatic health conditions, the associations seem to be more clearly confined to specific conditions such as epilepsy, heart disease, type 1 diabetes, or leukemia and central nervous system tumors, whereas adolescents experiencing other types of childhood cancers, asthma, or inflammatory bowel disease have often shown similar long-term educational outcomes as their peers without such health problems (Fletcher and Richards 2012; Persson et al. 2013; Mikkonen et al. 2018; Singh et al. 2015; Champaloux and Young 2015; Ghaderi et al. 2016; Lancashire et al. 2010).

The mechanisms of health-related selection to education are still poorly understood. We suggest that a theoretically important and policy-relevant distinction can be drawn between a process where health problems impair one's school performance, thus reducing the number of available options, and a process where health problems shape one's educational preferences regardless of their effects on school performance. This distinction is reminiscent of the one frequently made between primary and secondary effects of parental social class on educational attainment, originally introduced by Boudon (1974) and adopted by sociologists examining social stratification (Michelle Jackson et al. 2007; Erikson and Rudolphi 2010). Stated briefly, the idea is that parental social class affects both children's school performance (primary effect) and the educational choices children make within the range enabled by their past performance (secondary effect) (Michelle Jackson et al. 2007).

Along the same lines, we first argue that experiencing early-adolescent health problems may exhibit a primary effect on educational attainment by reducing school performance, measured by grade point average (GPA) in our application, and thus affecting actual opportunities for pursuing further education (primary effect). Part of this could be due to impaired academic ability

and functioning because previous research suggests poor early health to delay cognitive development (Bhutta et al. 2002; Boardman et al. 2002) and increase school absenteeism (Needham 2009; Haas and Fosse 2008; Hale and Viner 2018). On the other hand, because GPA is also a function of sustained effort, motivation, classroom behaviors, and teachers' expectations (McLeod, Uemura, and Rohrman 2012; Erikson and Rudolphi 2010), it can be seen to capture all activity that is either officially or unofficially appreciated at schools. When encountering an adolescent with emotional or behavioral disorders, teachers, parents and peers may come to reduce their expectations and withdraw their intellectual and social support as a reaction to increased negative interactions (Roeser, Eccles, and Strobel 1998). Such adolescents may eventually lower the expectations they hold for their own school performance and become disengaged from their peers and the school environment (McLeod and Fettes 2007; McLeod, Uemura, and Rohrman 2012; Needham, Crosnoe, and Muller 2004). These expectations could gradually become independent of individual teachers and current health status, as they are consolidated in the form of cumulative school records (Entwisle and Hayduk 1988).

Previous studies have identified teacher biases in grading by comparing teacher-rated grades with anonymously rated exam grades (Evensen 2019; Jæger and Møllegaard 2017). Evensen (2019) discovered that adolescents with attention or conduct problems received systematically lower teacher-rated grades than could have been expected based on their achievement in anonymously rated exams. In addition to such "behavior penalties" in grading, we suggest an even subtler process where the reductions in school performance among adolescents with health problems could be interpreted in a particularly negative manner—by the adolescents themselves or parents, teachers and school counselors advising them—and therefore lead to larger-than-expected changes in subsequent educational pathways. In other words, not only the grading itself but also the consequences of lower grades could vary between individuals with and without health problems.

Second, we posit that facing health problems during adolescence may reduce an adolescent's self-efficacy and experienced readiness to engage in further education, which could

produce less far-reaching or academically demanding educational choices even when school performance remains ostensibly intact (secondary effect). It seems plausible that reduced expectations among teachers, parents, or the adolescents themselves could operate, to some extent, independent of reductions in school performance because educational decisions can be seen as long-term investments that involve cost-benefit evaluation and only partially reflect past academic achievement (Michelle Jackson et al. 2007). Likewise, it has been suggested that experiencing health problems may alter an adolescent's future orientation and time-discounting practices such that he or she assigns more weight to the maximization of short-term gains than to long-term investments on human capital when making educational decisions (Haas, Glymour, and Berkman 2011). From an economic perspective, health problems reduce the expected utility of completing a longer education because they lower an individual's subjective life expectancy (Becker and Mulligan 1997).

A handful of previous studies have examined school performance as a mediator of the association between health problems and educational attainment. In one study, weakened academic performance explained the majority of the effect of self-rated health on timely high school graduation and around half of its effect on college attendance (Margot Jackson 2009). Another study showed that the association between self-reported psychological health complaints and entering tertiary education was fully accounted for by grade point average; for psychosomatic complaints, GPA was of less importance (Låftman and Magnusson 2017), whereas in a Norwegian study GPA covered more than two-thirds of the association between externalizing problems and non-completion of upper-secondary school and, among girls, 41% of the association of internalizing problems (Sagatun et al. 2014). McLeod et al. (2012) observed that the mutually adjusted coefficients of depression, attention problems, delinquency, and substance use with highest degree received were reduced by about half when controlling for high school GPA.

These investigations rely on self-reported health measures and none of them simultaneously assesses the impact of somatic and mental health problems, although the significance of weakened

school performance could vary according to the type of health problems. Despite being reliable and predictive of "objective" health status (Idler and Benyamini 1997), self-rated health may be less sensitive to changes in health status in early adolescence (Boardman 2006). Moreover, focus on a single binary outcome may conceal decisive horizontal stratification in educational pathways, since the same years of schooling in different study tracks may yield varying prospects for further education, workplace autonomy, and income development (Gerber and Cheung 2008). Finally, earlier analyses adopt the traditional mediation framework where health problems and school performance are not allowed to interact with each other (VanderWeele 2013b). Thus, they have been unable to examine the hypothesis that the consequences of lower grades could be more negative for adolescents with health problems, although this could be an important reinforcing component of the indirect pathway via weakened school performance.

1.2 Context and objectives

Finland provides an ideal setting for the study of health-related selection to education given that education is free-of-charge at all levels and that acceptance to upper-secondary education is, with few exceptions, determined by the combination of average grades at the end of compulsory school and one's preferred study program (Finnish National Agency for Education 2018a). Virtually all Finnish children (>99.5%) receive a certificate from compulsory basic education, which lasts nine years (with a voluntary tenth grade) and contains no tracking and only few optional courses (OECD 2013). After compulsory education, adolescents may apply to upper-secondary education, which lasts 2–4 years and divides into a general, academically oriented track and a vocational track (i.e., vocational education and training that prepares adolescents to work in manual or lower non-manual jobs). The two tracks are sometimes but rarely combined (Finnish National Agency for Education 2018a). Although the Finnish education system contains no absolute dead ends, only around 10% of students at universities providing academic higher education (master's degrees) have previously completed vocational upper-secondary education (Finnish National

Agency for Education 2018b). A similar kind of division between general and vocational secondary programs is commonplace across the OECD countries (OECD 2018a, 182).

Our study employs the counterfactual mediation framework (VanderWeele 2013b) to conceptually separate the direct (unexplained by school performance) and indirect (via school performance) effects of health problems on the likelihood of embarking on alternative upper-secondary educational pathways. The indirect effect via weakened school performance may be further divided into two parts (VanderWeele 2013b). In addition to the process where health problems first weaken school performance and consequently affect educational pathways (pure mediation), adolescents with health problems (or their parents and school counselors) could interpret the same reductions in school performance more negatively than their peers would when making educational choices (mediated interaction).

We focus on two upper-secondary educational outcomes: (1) completing some education versus completing no education and (2) for those completing some education, the choice between the general track and the vocational track. We hypothesize that encountering serious somatic conditions or mental disorders during early adolescence is associated with a higher likelihood of completing no upper-secondary education and, to a slightly smaller extent, a higher likelihood of completing the academically less-demanding vocational track instead of the general track. Furthermore, based on the small amount of available evidence, we hypothesize that both of these associations are mainly, but not totally, explained by weakened school performance. Without a useful reference point, we do not propose hypotheses on the relative importance of pure mediation and mediated interaction, but we expect to find evidence for both types of indirect effects. Figure 1 summarizes our causal assumptions.

Table 1 depicts our target effects and their definitions (adopted from Wang and Arah 2015). It should be noted that we refer to *effects* with their established names while being aware that our observational data only warrants the estimation of *associations*. The *total effect* summarizes the direct and indirect effects of health problems on education and consists of a *pure direct effect*, *pure*

indirect effect and *mediated interaction effect* in the three-way decomposition. The pure direct effect measures the effect of health problems on education through pathways other than school performance (mediator is held fixed at the value it would naturally take under no exposure), while the pure indirect effect measures the effect of health problems purely mediated by school performance, not accounting for the possible interaction between them (differential exposure). In both cases, the term "pure" refers to the fact that the effects have been dissociated from the mediated interaction, which accounts for the reciprocal part where health problems cause lower school performance and where the effect of weakened school performance is amplified by the presence of health problems (differential effect). The pure indirect effect and mediated interaction effect together form the indirect part of the total effect and can be used to calculate the *proportion mediated* if the effects are positive (VanderWeele 2013b).

We supplement these analyses with a policy-relevant examination of whether providing a certain level of school performance for all pupils would eradicate the direct effects of health problems. This is done by estimating *controlled direct effects* in a hypothetical situation that fixes school performance at different above-median levels for everyone. Based on the difference between the total effect and the controlled direct effect, we calculate the *proportion* (of the total effect) *eliminated* by this simulated hypothetical intervention (VanderWeele 2013a).

If we return to the above-mentioned theoretical distinction between primary and secondary effects, the pure indirect effect and mediated interaction effect arguably capture the primary effect due to weakened school performance, whereas the pure direct effect captures the secondary effect operating above and beyond school performance. Controlled direct effects provide a further evaluation of the secondary effects.

2 Methods

2.1 Dataset

All Finnish residents are assigned a personal identification number, which allows the individual-level linkage of information from different national registers. In this study, we had access to a 20% random sample (permission TK-53-525-11) of children who were born between 1987 and 1993 and belonged to the Finnish household population at the end of year 2000 (n=77,452). Statistics Finland drew the original sample, conducted the individual-level linkage of different data sources and delivered the data file in anonymized form.

Because we measured health problems at ages 10–16 and educational attainment at age 23, we excluded persons who were not alive or did not reside in mainland Finland at ages 10–16 or at age 23 (n=1,603). We measured parental socioeconomic factors at ages 10-14, and we also excluded those persons not dwelling with at least one parent at these ages (n=1,540). Furthermore, we excluded persons who had received treatment for intellectual disabilities (International Statistical Classification of Diseases and Related Health Problems, 10th revision, ICD-10 codes F70–F79) or pervasive developmental disorders (F84) by age 16 (n=963), since these adolescents less often attend regular upper-secondary schools. Finally, because the availability of grade point average—our measure of school performance at the end of basic schooling at age 16—was obtained from applications to upper-secondary education, we excluded persons who had not applied to upper-secondary education by age 20 (n=1,877, 2.4%, overall and n=873, 1.2%, following other exclusions). Our final study sample comprised 73,072 persons when analyzing the likelihood of not completing any upper-secondary education and 64,462 persons when analyzing the likelihood of completing the vocational track instead of the general track among those who completed some upper-secondary education by age 23.

2.2 Key measures

As outcome measures, we identified upper-secondary non-completion (0=some upper-secondary degree, 1=no education) and track choice at age 23 (0=general track or double degree, 1=vocational) using information from the Register of Degrees and Examinations maintained by Statistics Finland.

We assessed health problems at ages 10–16 using data provided by the National Institute for Health and Welfare on visits to inpatient and outpatient specialized care. Each record included a primary ICD-10 code for diagnosis. Our main analyses were based on broad groups of somatic conditions and mental disorders (see Supplementary Table 1 for precise definitions), defined as having at least one visit with a related diagnosis at ages 10-16. The group of somatic conditions effectively includes all types of conditions, such as asthma, epilepsy, type 1 diabetes, cancer, and severe infections (injuries, pregnancy-related diagnoses, and undefined symptoms excluded). We expect our use of a seven-year age span for the measurement of health problems to reduce both systematic and random variation in their detection.

For somatic conditions, we also created two indicators that capture more severe health problems: (1) having received inpatient or outpatient treatment for somatic health problems in four or more years during ages 10–16 and (2) pediatric complex chronic conditions. The latter category is based on a reviewed list of somatic conditions that are expected to last at least 12 months and require specialty care and possibly hospitalization (Feudtner et al. 2014). Within mental disorders, we chose two groups for further examination: depression or anxiety (internalizing) and attention deficit hyperactive disorder or conduct disorder (externalizing). These groups were chosen because they represent the most common types of mental disorders in early adolescence and because there has been discussion on the independence of their effects. To account for comorbidity, we adjusted for externalizing when estimating the effects of internalizing, and vice versa.

We measured school performance with grade point average (GPA), which is collected in the Joint Application Register of the Finnish National Board of Education. GPA is the arithmetic

mean of 11–13 theoretical subjects (graded from 4 to 10) on the compulsory school completion certificate, and it either fully or mostly defines one's chances of becoming accepted to the desired upper-secondary school, depending on the field of choice (when applying to vocational education, artistic and practical subjects are also considered). The final grading of pupils follows nationally determined learning goals, but the grading is not based on standardized tests. While this could introduce some school-level systematic variation in grading, it seems unlikely that such variation would be systematically associated with health problems and thus bias our results. GPA is recorded at the precision of 2 decimals and we treated it as a continuous variable.

2.3 Covariates

We controlled for a host of individual and family sociodemographic factors that could be simultaneously associated with adolescent health problems (or receiving treatment for them), school performance and educational pathways based on previous studies and theoretical considerations. The prevalence of treatment for mental disorders has increased among Finnish adolescents (Gyllenberg et al. 2018), and the popularity of the general and vocational tracks has varied over the years; thus, we included birth year dummies in all statistical models. Furthermore, we controlled for sex because not completing upper-secondary school is more common among men (Finnish National Agency for Education 2018b), and boys and girls show different patterns of health problems in adolescence especially with regard to mental health (Gyllenberg et al. 2018). We also adjusted for mother tongue (Finnish / Swedish / other), country of birth (Finland / other), region of residence at ages 10–14 (n=18), and urban–rural classification (7-point scale ranging from inner-urban to sparsely populated rural areas), which may capture differences in access to healthcare and upper-secondary education. Both poor adolescent health and low educational attainment are more prevalent in socially disadvantaged families (Viner et al. 2012; Sirin 2005); adjustments for family background included highest parental education (higher tertiary / lower tertiary / short-cycle tertiary / upper secondary / basic), five-year average household disposable income, most typical (mode)

family type (two parents / single parent) during ages 10–14. We also adjusted for the number of children under 18 living in the same household (sibship size). Finally, we adjusted for (continuous) maternal age at birth because the children of very young mothers have shown worse psychosocial and educational outcomes in previous studies (Fergusson and Woodward 1999).

2.4 Analytical procedure

To estimate the effects of interest, we used the parametric g-computation algorithm, the utility of which has been recently demonstrated in causal mediation settings (Wang and Arah 2015). G-computation can be considered a generalization of direct standardization and an approach to impute missing potential outcomes in the spirit of the counterfactual mediation framework (Westreich et al. 2015). As opposed to the traditional Baron and Kenny's method for mediation, g-computation avoids the issue of non-collapsibility, which hinders the comparison of coefficients between nested non-linear models. Compared with the alternative weighting-based approaches, parametric g-computation produces smaller standard errors and more stable estimates (Vansteelandt, Bekaert, and Lange 2012), and it also differs from the popular Karlson–Holm–Breen method by allowing for exposure–mediator interactions. Since detailed and accessible descriptions of the method in both mediation and non-mediation settings can be found elsewhere (Snowden, Rose, and Mortimer 2011; Wang and Arah 2015; Westreich et al. 2015), we provide only a brief summary of our empirical application. The procedure was similar for each exposure–outcome pair examined.

We first fitted an ordinary least squares model to estimate $E(M \mid X, Z)$ for the continuous mediator (grade point average) and a logistic regression model to estimate $logit[E(Y \mid X, M, Z)]$ for the binary outcome (non-completion / vocational track). For the continuous confounders maternal age, income, and sibship size, we chose a linear, quadratic or cubic polynomial depending on the p-value of the highest term ($p<0.1$) and, on similar grounds, decided to model the association between the mediator and the outcome as a third-degree polynomial with exposure–mediator interaction.

Using the parameters of the mediator model, we simulated two potential values of the mediator by setting x=1 (exposure to health problem) and x=0 (no exposure) for every individual in the sample. To reconstruct the original distribution of GPA, we added random variation from the normal distribution with variance based on the residual variance of the mediator model. Next, we simulated four potential outcomes (or their estimated probabilities) under x=1 and x=0 for everyone, given the parameters of the outcome model and the two simulated values of the mediator. For the purpose of calculating controlled direct effects, we further simulated potential outcomes under x=1 and x=0 while fixing GPA for everyone at two above-median values of interest: 8.50 and 9.09 (i.e., the 75^{th} and 90^{th} percentile of the GPA distribution). Finally, using the population-averaged values of the simulated potential outcomes, we calculated the relevant contrasts (risk differences) as presented in Table 1. We repeated the whole procedure 1000 times by resampling the original data to calculate bootstrapped 95% confidence intervals at the 2.5^{th} and 97.5^{th} percentiles of the estimate distribution.

Because our data included all household members, we were also able to conduct the three-way decomposition within sibships to adjust for all unobserved family-level confounding. Overall, there were 35,892 persons with at least one sibling in our study sample. We used fixed-effects linear probability models (adjusted for sex) because they allow for the extraction of sibship-level intercepts in an uncomplicated manner. Given that linear probability models could produce predicted probabilities that are smaller than 0 or larger than 1, we restricted the relatively few such cases to have values of 0 and 1, respectively (Bijlsma et al. 2019). We included sibship-level intercepts in the prediction, but the analytical procedure was otherwise similar to the one described above. We conducted all analyses with R version 3.5.1.

3 Results

Table 2 presents the distribution of the study population, mean grade point averages, and the distribution of different educational attainments at age 23 according to background factors.

Completing no upper-secondary education by age 23 was more common among adolescents with either somatic conditions or mental disorders, but the differences were larger with regard to mental disorders. Completing the general track was also much rarer among those with early-adolescent mental disorders, but differences in the proportion completing the vocational track were less evident. This illuminates the idea that a more nuanced view of the consequences of poor early health on education is achieved by focusing on the two choice situations separately.

Table 3 shows the decompositions of the effects of somatic conditions and mental disorders on not completing upper-secondary education. Estimated through the g-computation procedure and adjusted for all observed confounders, mental disorders were associated with a 19.2 percentage point increase in the probability of not completing upper-secondary education, whereas somatic conditions showed a much smaller 2.3 percentage point increase. In both cases, school performance mediated less than a third (31% for mental disorders and 22% for somatic conditions) of the total effects. The indirect part was mostly due to pure mediation (pure indirect effect) where the reductions in school performance lead to expectedly large changes in the subsequent educational pathways. However, for mental disorders there was also evidence of a small mediated interaction effect where health problems were not only associated with GPA but also augmented the negative association between GPA and non-completion. When we fixed GPA at the 75th and 90th percentile of the distribution for everyone to calculate controlled direct effects, the effect of mental disorders on non-completion was smaller but still evident, whereas the effect of somatic conditions virtually vanished at the 90th percentile. Table 3 also includes the results of the three-way decomposition conducted within sibships. The total effects of somatic conditions and mental disorders were around a half of the original effects, but the proportions mediated by GPA were only slightly smaller among siblings.

Table 4 presents similar decompositions with regard to choosing the vocational track instead of the general track. Mental disorders increased the probability of the vocational track by 16.0 percentage points and somatic conditions by 3.0 percentage points. School performance

mediated around half of these associations, a larger part than in the case of non-completion, and there was no evidence of mediated interaction for either somatic conditions or mental disorders. When GPA was fixed at the 90th percentile for everyone, mental disorders still increased the likelihood of vocational education by 5.8 percentage points, but the association of somatic conditions disappeared. Estimated within sibships, the total effects of health problems were up to two thirds weaker; the contribution of GPA remained similar for mental disorders but became slightly smaller for somatic conditions.

Table 5 presents the three-way decomposition for the more strictly defined health problem indicators. Notably, when exposure was defined as having received inpatient or outpatient treatment for somatic health problems in four or more years during ages 10–16, the effects were similar to the ones shown for somatic conditions in the main analysis. For complex chronic conditions, the total effects were roughly similar, but there was no evidence of an indirect effect via school performance. When estimating the effects of internalizing (depression or anxiety) and externalizing (attention deficit hyperactivity disorder or conduct disorder), we simultaneously included both types of conditions in the model to assess their independent effects. Whereas the effects of internalizing were comparable to the effects of the broader group of mental disorders, externalizing increased both non-completion and choosing the vocational track by around 30 percentage points. Moreover, externalizing showed slightly larger mediation proportions, and there was evidence of a small negative mediated interaction.

As a sensitivity analysis, we re-estimated the results separately for girls and boys (not shown). Somatic conditions increased non-completion by 3.1 (95% confidence interval: 2.5, 3.8) percentage points among girls and 1.3 (0.5, 2.0) percentage points among boys, whereas the probability of choosing the vocational track increased by 3.7 (2.5, 4.8) percentage points among girls and 0.7 (-0.4, 1.8) percentage points among boys. Proportions mediated by GPA were also slightly larger among girls. The effects of internalizing and mental disorders in general were similar regardless of sex. However, externalizing disorders exhibited slightly larger total effects on girls

than on boys with regard to both non-completion [33.6 (28.7, 38.0) vs. 27.3 (24.5, 30.6) percentage points, respectively] and choosing the vocational track [38.5 (33.9, 43.9) vs. 27.0 (23.8, 30.8) percentage points]. These differences were fully attributable to a stronger direct effect, unexplained by GPA, among girls.

The results on track choice only pertain to young people who completed upper-secondary education. Moreover, increases in the probability of completing the vocational track might not only reflect deliberate educational choices but could also indicate failure to complete the general track. To evaluate whether our results can be interpreted in terms of actual decisions, we used data on the national application process that was available for adolescents born in 1992-1993 (n=24,175). Supplementary Table 2 reports an analysis on the association between health problems and choosing the vocational track as the first preference in the joint application process. Although this analysis was independent of both enrollment and completion, results on first preference were identical with the ones based on attained education. Data on the application process also enabled us to assess the most critical consequence of weakened school performance: failure to enter the preferred study program. To estimate its contribution to the association between health problems and non-completion, Supplementary Table 3 presents an analysis where failure to enter the preferred study program (prevalence: 20.7%) substituted for GPA as the mediator. In the case of both somatic conditions and mental disorders, it mediated around 10% of the total effect of health problems on non-completion.

4 Discussion

Using a large register-based dataset, we show that early-adolescent health problems are significant predictors of both non-completion and track choice in upper-secondary education and that these associations are only partially explained by weakened school performance at the end of compulsory education. We observe that while the weight of mental disorders in shaping both

outcomes is especially noteworthy, even somatic conditions exhibit small but robust increases in the likelihood of completing no upper-secondary education and choosing the less academically oriented vocational track. At the same time, the results imply that weakened school performance accounts for a smaller part of the differences in non-completion (one third at most) than in track choice (around half) between adolescents with and without health problems. The same patterns can be observed within sibships, albeit on a smaller scale. Furthermore, according to our simulated hypothetical intervention, mental disorders might still be associated with educational pathways even if all adolescents had excellent school performance.

Taking advantage of the fairly recently introduced three-way decomposition in mediation analysis (VanderWeele 2013b), our study is the first to elucidate that the indirect association is mostly due to pure mediation by school performance instead of mediated interaction. This supports the interpretation that health problems first reduce school performance by a certain amount and educational choices are then affected as much as could be expected for such a reduction in school performance—no more, no less. Only for the association between mental disorders and non-completion do we observe evidence of a small mediated interaction, possibly signaling a process whereby teachers, parents, school counselors or the adolescents themselves interpret the same level of school performance more negatively than their peers without a history of treatment for mental disorders. Previous research has identified teacher biases in the grading of adolescents with attention or conduct problems (Evensen 2019), but our results suggest a further "bias" where lower grades increase educational non-completion more than could be expected among those experiencing health problems. Such process could possibly be explained by stigmatization and increased negative interactions in the school environment (McLeod and Kaiser 2004), but examining this question would require more detailed data.

In our analysis, the indirect path via school performance accounted for less than a third of the differences in non-completion, but previous studies addressing the same question have shown mediation proportions of 50% or more in the US (Margot Jackson 2009) and Norway (Sagatun et al.

2014). One study from Sweden, however, found GPA important in explaining the effects of psychological health complaints, but not those of psychosomatic complaints (Låftman and Magnusson 2017). Part of these differences could reflect our data on specialized care, capturing more severe cases on average, whereas earlier studies have measured either self-rated health or self-reported health complaints. Mediation proportions could also be underestimated in our analysis if a higher GPA is an implicit sign of less severe problems within the broadly defined groups of health problems. However, the fact that the mediation proportions were similar for the more strictly defined exposures mitigated such concerns.

We are unaware of other studies that have examined the link stretching from health problems to school performance and subsequently to track choice in a similar manner. One previous study collapsed all degrees into a single ordinal outcome and controlled for GPA (McLeod, Uemura, and Rohrman 2012), but this approach conceals horizontal stratification. Moreover, considering that the strength of health selection may vary according to life stage (Kane et al. 2018), treating education as an ordinal or continuous variable could also obscure health-related vertical stratification. Overall, previous studies have only analyzed track choice when comparing the predictive power of health problems for dropout between different educational tracks or levels (Sagatun et al. 2014; Hjorth et al. 2016). Although one need not consider the general track a more commendable alternative than the vocational track, the observed differences in track choice could reflect structural, potentially modifiable obstacles that young people with health problems encounter when navigating the education system.

Crucially, our supplementary analysis, which was based on school applications, aligned with our main analyses that examined completed education. **This supports the interpretation that adolescents with health problems more often *choose* the vocational track, but it remains unclear whether this is due to altered preferences, anticipatory effects (expecting not to be able to enter the general track), or reduced educational expectations held by significant others. Regardless of the origin of educational decisions, being able to execute them appears**

beneficial. In our further supplementary analysis, failure to enter the desired study program—the immediate consequence of poor school performance—explained one tenth of the association between health problems and non-completion. Despite indicating a smaller indirect effect than the one shown for GPA, the fact that it is based on first-time applicants, who have an opportunity to take an additional study year to raise their grades and whose preferences may change, renders it noteworthy.

Different education systems pose a challenge to the generalizability of health selection research (Suhrcke and de Paz Nieves 2011). The weight of our results is emphasized by the fact that they are obtained in a universalistic education and welfare system, which could be expected to bridge gaps in access to healthcare and education and thus reduce health-related selection to education. Our study is in line with earlier research from the UK (Goodman, Joyce, and Smith 2011), the Netherlands (Uiters et al. 2014), Canada (Currie et al. 2010), and Norway (De Ridder et al. 2013), demonstrating that early somatic health problems are on average rather weakly associated with adult social outcomes. On the other hand, some specific somatic conditions are likely to show stronger associations, as discovered by investigations from Finland (Mikkonen et al. 2018) and the US (Champaloux and Young 2015) simultaneously measuring the effects of several specific conditions. Overall, it is surprising that increased length (treatment in four or more years) or severity (complex chronic conditions) of exposure does not strengthen the observed associations between somatic health problems and educational outcomes compared with the broader category. One plausible explanation is that children with more serious health conditions may more often attend hospital schools or receive other kinds of special support, which promotes their schooling and which we cannot observe with register data. Moreover, the average associations may contain masked heterogeneity because for some adolescents a chronic illness might act as a reason to pursue an academic degree that leads to a physically non-demanding white-collar profession (Teachman 2012). Finally, it should be noted that in the case of life-threatening health conditions, only those individuals who survived until age 23 contributed to our results.

Similar to most previous studies, we observe robust effects for mental disorders even when the estimation is based on within-sibship variation (Fletcher 2010; Evensen et al. 2016). In line with earlier research, we also show stronger effects for externalizing than internalizing disorders (Miech et al. 2002; McLeod and Kaiser 2004). However, our observation that internalizing disorders are negatively associated with educational outcomes regardless of co-occurring externalizing disorders is in stark contrast to the study of Evensen et al. (2016) from Norway where the effects of internalizing disappeared following adjustment for attention and conduct problems. Even so, our study does not necessarily challenge their results because they measured mental disorders based on self-reported symptom questionnaires, whereas we defined comorbidity based on a stricter measure of having received diagnoses related to both internalizing and externalizing in specialized care.

Our study evaded the typical problems of survey studies related to recall bias, preferential reporting, non-response, attrition, and large standard errors. Moreover, with register data we were able to distinguish the temporal ordering of health problems and educational outcomes more precisely than would have been possible with cross-sectional or retrospective data, and it also made the identification of different types of health problems more objective than in previous mediation studies. However, it is essential to note that we were only able to measure health problems that received inpatient or outpatient care at least for a short while. While this enabled us to focus on more severe health conditions than in most previous research, it also left out high-risk health behaviors and other lifestyle-related factors (e.g. smoking, obesity, and physical activity) as well as recurrent minor illnesses, which could prove important when evaluating the general contribution of adolescent health to educational outcomes.

Especially in the case of mental disorders, the reliance on care records could also bias our results if unobserved factors affect the likelihood of seeking or receiving treatment and some untreated persons with mental disorders have been misclassified as not having mental disorders. More specifically, this may have caused us to underestimate the total effect but is unlikely to have biased our decomposition to direct and mediated effects. On the positive side, we were able to

control for the most evident socioeconomic and regional correlates of treatment seeking, which is likely to render any remaining biases small. Gaps in treatment are further narrowed by the fact that pediatric specialized care is highly subsidized and provided by the public sector in Finland (Ministry of Social Affairs and Health 2013). As another point of concern, weakened school performance, an important indicator of child and adolescent functioning, could promote the identification of emotional and behavioral disorders in the school environment. In such cases, the indirect effects of mental disorders could be partially biased. However, Finnish pupils attend three extensive health checks during compulsory education, which curtails the role of teachers or parents in the detection of problems (Ministry of Social Affairs and Health 2013).

We did not have information on GPA for 1.2% of the study population. To evaluate the impact of this exclusion, we reproduced the four total effects reported in Table 3 and Table 4 by including those with a missing GPA (not shown). Although health problems were more common in this group, the total effects became only 2-4 percent larger than in our principal study sample. Thus, this exclusion criterion made our results only slightly more conservative.

The g-computation procedure produces stable and efficient estimates but relies on the correct specification of both the mediator and the outcome model when using observational data. In addition to our broad adjustment for exposure–outcome, exposure–mediator and mediator–exposure confounders related to sociodemographic background, we were able to use sibling fixed-effects models to further adjust for all unobserved factors shared within a family. These models also partially account for cognitive and non-cognitive skills and other common, possibly genetic, factors behind health and education (Boardman, Domingue, and Daw 2015; Duke and Macmillan 2016). Nevertheless, our sibling comparison results should be interpreted with care because the reliance on within-family variation in health problems could hamper their generalizability. Potential spillover effects running from health problems to siblings' educational outcomes could also render their results too conservative. A further assumption with our three-way effect decomposition is the absence of exposure-induced mediator–outcome confounders (VanderWeele 2013b), but we are

unaware of any previous empirical evidence or theoretical consideration on confounders between GPA and educational choices that are influenced by health problems.

The pathways that we describe in this study are best understood against long-term life course processes. Health problems show strong path dependency in adolescence and young adulthood (Kane et al. 2018), and it appears probable that the effects we observe for health problems at age 10-16 are partially due to the persistence and perpetuation of these problems. Correspondingly, although our study focused on health problems in early adolescence, some of them may have emerged at an earlier age and affected early schooling in ways not captured by our study. To analyze the contribution of early-adolescent health independent of earlier and later selection processes, future studies would benefit from a dataset with longitudinal measurement of both health problems and school functioning that begins at an early age and extends until the end of upper-secondary education. Such data would also enable them to examine the potential reciprocity of the associations between health problems and weak school performance.

Regardless of the exact timing of the processes at work, our results provide compelling take-home messages for life-course research on health inequalities and social stratification. Although failure to complete any upper-secondary education is likely to have the largest negative impact on individual life chances, it is noteworthy that adolescents with health problems are also more likely to choose the vocational track, which rarely leads to the highest-paying, more autonomous professions. Such subtle health-related differences in early educational decisions are more likely neglected than differences in non-completion and may contribute to socioeconomic health disparities in the long-term (Palloni 2006). Moreover, although both poor adolescent health and adverse educational outcomes are more prevalent in socially disadvantaged families (Sirin 2005; Viner et al. 2012), our results evince that health selection cannot be reduced to disadvantaged social background alone.

Reminiscent of the theoretical distinction between primary and secondary effects of parental social class on offspring education (Boudon 1974), our study separates the primary effect

of health problems due to weakened school performance from the secondary effect of health problems that operates above and beyond school performance. We show that health problems have a non-negligible primary effect that is mainly due to a process where health-induced reductions in GPA lead to differences in educational pathways that are as large as could be expected for such reductions. This finding encourages future research and policy efforts to target factors that disturb the schooling of adolescents with health problems to the extent that their formal school performance is impaired. However, considering that GPA reflects not only academic aptitude but also sustained effort and teacher–student interactions (McLeod, Uemura, and Rohrman 2012; Erikson and Rudolphi 2010), interventions that only aim at promoting academic skills could prove inadequate in reducing health-related selection. Previous research on adolescents with mental disorders emphasizes the importance of negative expectations and social responses as explanations for disrupted educational trajectories (McLeod and Fettes 2007).

Simultaneously, our investigation shows that a significant part of the associations cannot be ascribed to reductions in school performance. This secondary effect, as we call it, ultimately remains a black box in our analysis. However, its presence highlights that GPA cannot be considered an all-encompassing predictor of sustained educational success among adolescents with health problems, which can be seen as a deviation from the general path-dependency of human capital formation (Kane et al. 2018). Post-compulsory education requires increasing independence from young people, which can disclose underlying problems. On the other hand, it has been suggested that health problems may reduce future orientation towards time-consuming investments, such as higher education (Becker and Mulligan 1997; Haas, Glymour, and Berkman 2011). In our analysis, the increases observed in vocational education could be interpreted as signals of such processes at work, but this and other plausible mechanisms explaining the secondary effect require further investigation. In any event, it seems that adolescents with health problems, especially mental disorders, should be given special attention and support in their educational careers regardless of the formal school performance they demonstrate.

Becker, Gary S., and Casey B. Mulligan. 1997. “The Endogenous Determination of Time Preference.” *Quarterly Journal of Economics* 112(3): 729–58. https://doi.org/10.1162/003355397555334.

Bhutta, Adnan T., Mario A. Cleves, Patrick H. Casey, Mary M. Cradock, and Karthika J. S. Anand. 2002. “Cognitive and Behavioral Outcomes of School-Aged Children Who Were Born Preterm: A Meta-Analysis.” *Journal of the American Medical Association* 288(6): 728–37. https://doi.org/10.1001/jama.288.6.728.

Bijlsma, Maarten J., Ben Wilson, Lasse Tarkiainen, Mikko Myrskylä, and Pekka Martikainen. 2019. “The Impact of Unemployment on Antidepressant Purchasing: Adjusting for Unobserved Time-Constant Confounding in the g-Formula.” *Epidemiology* 30(3): 388–95. https://doi.org/10.1097/EDE.0000000000000985.

Boardman, Jason D. 2006. “Self-Rated Health among U.S. Adolescents.” *Journal of Adolescent Health* 38(4): 401–8. https://doi.org/10.1016/j.jadohealth.2005.01.006.

Boardman, Jason D., Benjamin W. Domingue, and Jonathan Daw. 2015. “What Can Genes Tell Us about the Relationship between Education and Health?” *Social Science and Medicine* 127: 171–80. https://doi.org/10.1016/j.socscimed.2014.08.001.

Boardman, Jason D., Daniel A. Powers, Yolanda C. Padilla, and Robert A. Hummer. 2002. “Low Birth Weight, Social Factors, and Developmental Outcomes among Children in the United States.” *Demography* 39(2): 353–68. https://doi.org/10.2307/3088343.

Boudon, Raymond. 1974. *Education, Opportunity, and Social Inequality*. New York: Wiley.

Brekke, Idunn. 2015. “Health and Educational Success in Adolescents: A Longitudinal Study.” *BMC Public Health* 15 (December): 619. https://doi.org/10.1186/s12889-015-1966-0.

Breslau, Joshua, Michael Lane, Nancy Sampson, and Ronald C. Kessler. 2008. “Mental Disorders and Subsequent Educational Attainment in a US National Sample.” *Journal of Psychiatric*

Research 42(9): 708–16. https://doi.org/10.1016/j.jpsychires.2008.01.016.

Champaloux, Steven W., and Deborah R. Young. 2015. "Childhood Chronic Health Conditions and Educational Attainment: A Social Ecological Approach." *Journal of Adolescent Health* 56(1): 98–105. https://doi.org/10.1016/j.jadohealth.2014.07.016.

Crump, Casey, Diana Rivera, Rebecca London, Melinda Landau, Bill Erlendson, and Eunice Rodriguez. 2013. "Chronic Health Conditions and School Performance among Children and Youth." *Annals of Epidemiology* 23(4): 179–84. https://doi.org/10.1016/j.annepidem.2013.01.001.

Currie, Janet, Mark Stabile, Phongsack Manivong, and Leslie L. Roos. 2010. "Child Health and Young Adult Outcomes." *Journal of Human Resources* 45(3): 517–48. https://doi.org/10.1353/jhr.2010.0013.

Duke, Naomi, and Ross Macmillan. 2016. "Schooling, Skills, and Self-Rated Health: A Test of Conventional Wisdom on the Relationship between Educational Attainment and Health." *Sociology of Education* 89(3): 171–206. https://doi.org/10.1177/0038040716653168.

Eide, Eric R., Mark H. Showalter, and Dan D. Goldhaber. 2010. "The Relation between Children's Health and Academic Achievement." *Children and Youth Services Review* 32(2): 231–38. https://doi.org/10.1016/j.childyouth.2009.08.019.

Entwisle, Doris R., and Leslie Alec Hayduk. 1988. "Lasting Effects of Elementary School." *Sociology of Education* 61(3): 147–59. https://doi.org/10.2307/2112624.

Erikson, Robert, and Frida Rudolphi. 2010. "Change in Social Selection to Upper Secondary School - Primary and Secondary Effects in Sweden." *European Sociological Review* 26(3): 291–305. https://doi.org/10.1093/esr/jcp022.

Esch, Pascale, Valéry Bocquet, Charles Pull, Sophie Couffignal, Torsten Lehnert, Marc Graas, Laurence Fond-Harmant, and Marc Ansseau. 2014. "The Downward Spiral of Mental

Disorders and Educational Attainment: A Systematic Review on Early School Leaving." *BMC Psychiatry* 14(237). https://doi.org/10.1186/s12888-014-0237-4.

Evensen, Miriam. 2019. "Adolescent Mental Health Problems, Behaviour Penalties, and Distributional Variation in Educational Achievement." *European Sociological Review* 35(4): 474–90. https://doi.org/10.1093/esr/jcz021.

Evensen, Miriam, Torkild Hovde Lyngstad, Ole Melkevik, and Arnstein Mykletun. 2016. "The Role of Internalizing and Externalizing Problems in Adolescence for Adult Educational Attainment: Evidence from Sibling Comparisons Using Data from the Young HUNT Study." *European Sociological Review* 32(5): 552–66. https://doi.org/10.1093/esr/jcw001.

Fergusson, David M., and Lianne J. Woodward. 1999. "Maternal Age and Educational and Psychosocial Outcomes in Early Adulthood." *Journal of Child Psychology and Psychiatry and Allied Disciplines* 40(3): 479–89. https://doi.org/10.1017/S0021963098003722.

———. 2002. "Mental Health, Educational, and Social Role Outcomes of Adolescents with Depression." *Archives of General Psychiatry* 59(3): 225–31. https://doi.org/10.1001/archpsyc.59.3.225.

Feudtner, Chris, James A. Feinstein, Wenjun Zhong, Matt Hall, and Dingwei Dai. 2014. "Pediatric Complex Chronic Conditions Classification System Version 2: Updated for ICD-10 and Complex Medical Technology Dependence and Transplantation." *BMC Pediatrics* 14(199). https://doi.org/10.1186/1471-2431-14-199.

Finnish National Agency for Education. 2018a. "Education System." 2018. https://www.oph.fi/english/education_system.

———. 2018b. "Vipunen. Education Statistics Finland." 2018. https://vipunen.fi/en-gb/.

Fletcher, Jason M. 2010. "Adolescent Depression and Educational Attainment: Results Using Sibling Fixed Effects." *Health Economics* 19(7): 855–71. https://doi.org/10.1002/hec.1526.

Fletcher, Jason M., and Michael R. Richards. 2012. “Diabetes’s ‘health Shock’ to Schooling and Earnings: Increased Dropout Rates and Lower Wages and Employment in Young Adults.” *Health Affairs* 31(1): 27–34. https://doi.org/10.1377/hlthaff.2011.0862.

Fried, Ronna, Carter Petty, Stephen V. Faraone, Laran L. Hyder, Helen Day, and Joseph Biederman. 2016. “Is ADHD a Risk Factor for High School Dropout? A Controlled Study.” *Journal of Attention Disorders* 20(5): 383–89. https://doi.org/10.1177/1087054712473180.

Frissen, Aleida, Ritsaert Lieverse, Machteld Marcelis, Marjan Drukker, and Philippe Delespaul. 2015. “Psychotic Disorder and Educational Achievement: A Family-Based Analysis.” *Social Psychiatry and Psychiatric Epidemiology* 50(10): 1511–18. https://doi.org/10.1007/s00127-015-1082-6.

Gerber, Theodore P., and Sin Yi Cheung. 2008. “Horizontal Stratification in Postsecondary Education: Forms, Explanations, and Implications.” *Annual Review of Sociology* 34(1): 299–318. https://doi.org/10.1146/annurev.soc.34.040507.134604.

Ghaderi, Sara, Anders Engeland, Maria Winther Gunnes, Dag Moster, Ellen Ruud, Astri Syse, Finn Wesenberg, and Tone Bjørge. 2016. “Educational Attainment among Long-Term Survivors of Cancer in Childhood and Adolescence: A Norwegian Population-Based Cohort Study.” *Journal of Cancer Survivorship* 10(1): 87–95. https://doi.org/10.1007/s11764-015-0453-z.

Goodman, Alissa, Robert Joyce, and James P. Smith. 2011. “The Long Shadow Cast by Childhood Physical and Mental Problems on Adult Life.” *Proceedings of the National Academy of Sciences* 108(15): 6032–37. https://doi.org/10.1073/pnas.1016970108.

Goulding, Sandra M., Victoria H. Chien, and Michael T. Compton. 2010. “Prevalence and Correlates of School Drop-out Prior to Initial Treatment of Nonaffective Psychosis: Further Evidence Suggesting a Need for Supported Education.” *Schizophrenia Research* 116(2–3): 228–33. https://doi.org/10.1016/j.schres.2009.09.006.

Gyllenberg, David, Mikko Marttila, Reijo Sund, Elina Jokiranta-Olkoniemi, André Sourander, Mika Gissler, and Tiina Ristikari. 2018. “Temporal Changes in the Incidence of Treated Psychiatric and Neurodevelopmental Disorders during Adolescence: An Analysis of Two National Finnish Birth Cohorts.” *The Lancet. Psychiatry* 5(3): 227–36. https://doi.org/10.1016/S2215-0366(18)30038-5.

Haas, Steven A., and Nathan Edward Fosse. 2008. “Health and the Educational Attainment of Adolescents: Evidence from the NLSY97.” *Journal of Health and Social Behavior* 49(2): 178–92. https://doi.org/10.1177/002214650804900205.

Haas, Steven A., M. Maria Glymour, and Lisa F. Berkman. 2011. “Childhood Health and Labor Market Inequality over the Life Course.” *Journal of Health and Social Behavior* 52(3): 298–313. https://doi.org/10.1177/0022146511410431.

Hale, Daniel R., Leonardo Bevilacqua, and Russell M. Viner. 2015. “Adolescent Health and Adult Education and Employment: A Systematic Review.” *Pediatrics* 136(1): 128–40. https://doi.org/10.1542/peds.2014-2105.

Hale, Daniel R., and Russell M. Viner. 2018. “How Adolescent Health Influences Education and Employment: Investigating Longitudinal Associations and Mechanisms.” *Journal of Epidemiology and Community Health* 72(6): 465–70. https://doi.org/10.1136/jech-2017-209605.

Hjorth, Cathrine F., Line Bilgrav, Louise Sjørslev Frandsen, Charlotte Overgaard, Christian Torp-Pedersen, Berit Nielsen, and Henrik Bøggild. 2016. “Mental Health and School Dropout across Educational Levels and Genders: A 4.8-Year Follow-up Study.” *BMC Public Health* 16(1). https://doi.org/10.1186/s12889-016-3622-8.

Huurre, Taina, Hillevi Aro, Ossi Rahkonen, and Erkki Komulainen. 2006. “Health, Lifestyle, Family and School Factors in Adolescence: Predicting Adult Educational Level.” *Educational*

Research 48(1): 41–53. https://doi.org/10.1080/00131880500498438.

Idler, Ellen L., and Yael Benyamini. 1997. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior* 38(1): 21–37. https://doi.org/10.2307/2955359.

Isohanni, Irene, Peter B. Jones, Marjo-Riitta Jarvelin, Pentti Nieminen, Paula Rantakallio, Jari Jokelainen, Tim J. Croudace, and Matti Isohanni. 2001. "Educational Consequences of Mental Disorders Treated in Hospital. A 31-Year Follow-up of the Northern Finland 1966 Birth Cohort." *Psychological Medicine* 31(2): 339–49. https://doi.org/10.1017/S003329170100304X.

Jackson, Margot. 2009. "Understanding Links Between Adolescent Health and Educational Attainment." *Demography* 46(4): 671–94. https://doi.org/10.1353/dem.0.0078.

Jackson, Michelle, Robert Erikson, John H. Goldthorpe, and Meir Yaish. 2007. "Primary and Secondary Effects in Class Differentials in Educational Attainment: The Transition to A-Level Courses in England and Wales." *Acta Sociologica* 50(3): 211–29. https://doi.org/10.1177/0001699307080926.

Jæger, Mads Meier, and Stine Møllegaard. 2017. "Cultural Capital, Teacher Bias, and Educational Success: New Evidence from Monozygotic Twins." *Social Science Research* 65: 130–44. https://doi.org/10.1016/j.ssresearch.2017.04.003.

Johnson, Sheri L., Carol A. Winett, Björn Meyer, William J. Greenhouse, and Ivan Miller. 1999. "Social Support and the Course of Bipolar Disorder." *Journal of Abnormal Psychology* 108(4): 558–66. https://doi.org/10.1037/0021-843X.108.4.558.

Kane, Jennifer B., Kathleen Mullan Harris, S. Philip Morgan, and David K. Guilkey. 2018. "Pathways of Health and Human Capital from Adolescence into Young Adulthood." *Social Forces* 96(3): 949–76. https://doi.org/10.1093/sf/sox079.

Kessler, Roland C., Cindy L. Foster, William B. Saunders, and Paul E. Stang. 1995. "Social Consequences of Psychiatric Disorders, I: Educational Attainment." *The American Journal of Psychiatry* 152(7): 1026–32. https://doi.org/10.1176/ajp.152.7.1026.

Låftman, Sara B., and Charlotta Magnusson. 2017. "Do Health Complaints in Adolescence Negatively Predict the Chance of Entering Tertiary Education in Young Adulthood?" *Scandinavian Journal of Public Health* 45(8): 878–85. https://doi.org/10.1177/1403494817713649.

Lancashire, Emma R., Clare Frobisher, Raoul C. Reulen, David L. Winter, Adam Glaser, and Michael M. Hawkins. 2010. "Educational Attainment among Adult Survivors of Childhood Cancer in Great Britain: A Population-Based Cohort Study." *Journal of the National Cancer Institute* 102(4): 254–70. https://doi.org/10.1093/jnci/djp498.

Layte, Richard, and Cathal McCrory. 2013. "Paediatric Chronic Illness and Educational Failure: The Role of Emotional and Behavioural Problems." *Social Psychiatry and Psychiatric Epidemiology* 48(8): 1307–16. https://doi.org/10.1007/s00127-012-0609-3.

Lê, Félice, Ana Diez Roux, and Hal Morgenstern. 2013. "Effects of Child and Adolescent Health on Educational Progress." *Social Science & Medicine* 76: 57–66. https://doi.org/http://dx.doi.org/10.1016/j.socscimed.2012.10.005.

McLeod, Jane D., and Danielle L. Fettes. 2007. "Trajectories of Failure: The Educational Careers of Children with Mental Health Problems." *American Journal of Sociology* 113(3): 653–701. https://doi.org/10.1086/521849.

McLeod, Jane D., and Karen Kaiser. 2004. "Childhood Emotional and Behavioral Problems and Educational Attainment." *American Sociological Review* 69(5): 636–58. https://doi.org/10.1177/000312240406900502.

McLeod, Jane D, Ryotaro Uemura, and Shawna Rohrman. 2012. "Adolescent Mental Health,

Behavior Problems, and Academic Achievement." *Journal of Health and Social Behavior* 53(4): 482–97. https://doi.org/10.1177/0022146512462888.

McMahon, Walter W., and Moses Oketch. 2013. "Education's Effects on Individual Life Chances and On Development: An Overview." *British Journal of Educational Studies* 61(1): 79–107. https://doi.org/10.1080/00071005.2012.756170.

Melkevik, Ole, Wendy Nilsen, Miriam Evensen, Anne Reneflot, and Arnstein Mykletun. 2016. "Internalizing Disorders as Risk Factors for Early School Leaving: A Systematic Review." *Adolescent Research Review* 1(3): 245–55. https://doi.org/10.1007/s40894-016-0024-1.

Miech, Richard A., Avshalom Caspi, Terrie E. Moffitt, Bradley R. Entner Wright, and Phil A. Silva. 2002. "Low Socioeconomic Status and Mental Disorders: A Longitudinal Study of Selection and Causation during Young Adulthood." *American Journal of Sociology* 104(4): 1096–131. https://doi.org/10.1086/210137.

Mikkonen, Janne, Heta Moustgaard, Hanna Remes, and Pekka Martikainen. 2018. "The Population Impact of Childhood Health Conditions on Dropout from Upper-Secondary Education." *Journal of Pediatrics* 196: 283-90.e4. https://doi.org/10.1016/j.jpeds.2018.01.034.

Ministry of Social Affairs and Health. 2013. "Health Care in Finland." Brochures 2eng. Helsinki: Ministry of Social Affairs and Health. http://urn.fi/URN:ISBN:978-952-00-3395-8.

Needham, Belinda L. 2009. "Adolescent Depressive Symptomatology and Young Adult Educational Attainment: An Examination of Gender Differences." *Journal of Adolescent Health* 45(2): 179–86. https://doi.org/10.1016/j.jadohealth.2008.12.015.

Needham, Belinda L., Robert Crosnoe, and Chandra Muller. 2004. "Academic Failure in Secondary School: The Inter-Related Role of Health Problems and Educational Context." *Social Problems* 51(4): 569–86. https://doi.org/10.1525/sp.2004.51.4.569.

OECD. 2004. "Completing the Foundation for Lifelong Learning: An OECD Survey of Upper

Secondary Schools." Paris: OECD Publishing. https://doi.org/10.1787/9789264103733-en.

———. 2013. "Education Policy Outlook: Finland." Paris. http://www.oecd.org/education/EDUCATION POLICY OUTLOOK FINLAND_EN.pdf.

———. 2018a. *Education at a Glance 2018: OECD Indicators*. Paris: OECD Publishing.

———. 2018b. "How Do Admission Systems Affect Enrolment in Public Tertiary Education?" 63. Education Indicators in Focus. Paris.

Palloni, Alberto. 2006. "Reproducing Inequalities: Luck, Wallets, and the Enduring Effects of Childhood Health." *Demography* 43(4): 587–615. https://doi.org/10.1353/dem.2006.0036.

Persson, Sofie, Gisela Dahlquist, Ulf-Göran Gerdtham, and Katarina Steen Carlsson. 2013. "Impact of Childhood-Onset Type 1 Diabetes on Schooling: A Population-Based Register Study." *Diabetologia* 56(6): 1254–62. https://doi.org/10.1007/s00125-013-2870-8.

Ridder, Karin A. A. De, Kristine Pape, Roar Johnsen, Turid Lingaas Holmen, Steinar Westin, and Johan Håkon Bjørngaard. 2013. "Adolescent Health and High School Dropout: A Prospective Cohort Study of 9000 Norwegian Adolescents (The Young-HUNT)." *PLoS ONE* 8(9). https://doi.org/10.1371/journal.pone.0074954.

Roeser, Robert W., Jacquelynne S. Eccles, and Karen R. Strobel. 1998. "Linking the Study of Schooling and Mental Health: Selected Issues and Empirical Illustrations at the Level of the Individual." *Educational Psychologist* 33(4): 153–76. https://doi.org/10.1207/s15326985ep3304_2.

Sagatun, Ase, Sonja Heyerdahl, Tore Wentzel-Larsen, and Lars Lien. 2014. "Mental Health Problems in the 10th Grade and Non-Completion of Upper Secondary School: The Mediating Role of Grades in a Population-Based Longitudinal Study." *BMC Public Health* 14(1): 16. https://doi.org/10.1186/1471-2458-14-16.

Singh, Harminder, Zoann Nugent, Marni Brownell, Laura E. Targownik, Leslie L. Roos, and

Charles N. Bernstein. 2015. “Academic Performance among Children with Inflammatory Bowel Disease: A Population-Based Study.” *The Journal of Pediatrics* 166(5): 1128–33. https://doi.org/10.1016/j.jpeds.2014.12.010.

Sirin, Selcuk R. 2005. “Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research.” *Review of Educational Research* 75(3): 417–53. https://doi.org/10.3102/00346543075003417.

Snowden, Jonathan M., Sherri Rose, and Kathleen M. Mortimer. 2011. “Implementation of G-Computation on a Simulated Data Set: Demonstration of a Causal Inference Technique.” *American Journal of Epidemiology* 173(7): 731–38. https://doi.org/10.1093/aje/kwq472.

Suhrcke, Marc, and Carmen de Paz Nieves. 2011. “The Impact of Health and Health Behaviours on Educational Outcomes in High-Income Countries: A Review of the Evidence.” Copenhagen.

Teachman, Jay. 2012. “Health Limitations and Post-Secondary School Enrollment.” *Population Research and Policy Review* 31(1): 85–96. https://doi.org/10.1007/s11113-011-9225-2.

Uiters, Ellen, Erica Maurits, Mariël Droomers, Marieke Zwaanswijk, Robert A Verheij, and Fons van der Lucht. 2014. “The Association between Adolescents’ Health and Disparities in School Career: A Longitudinal Cohort Study.” *BMC Public Health* 14(1): 1104. https://doi.org/10.1186/1471-2458-14-1104.

VanderWeele, Tyler J. 2013a. “Policy-Relevant Proportions for Direct Effects.” *Epidemiology (Cambridge, Mass.)* 24(1): 175–76. https://doi.org/10.1097/EDE.0b013e3182781410.

———. 2013b. “A Three-Way Decomposition of a Total Effect into Direct, Indirect, and Interactive Effects.” *Epidemiology (Cambridge, Mass.)* 24(2): 224–32. https://doi.org/10.1097/EDE.0b013e318281a64e.

Vansteelandt, Stijn, Maarten Bekaert, and Theis Lange. 2012. “Imputation Strategies for the Estimation of Natural Direct and Indirect Effects.” *Epidemiologic Methods* 1(1): 131–58.

https://doi.org/10.1515/2161-962X.1014.

Viner, Russell M., Elizabeth M. Ozer, Simon Denny, Michael Marmot, Michael Resnick, Adesegun Fatusi, and Candace Currie. 2012. "Adolescence and the Social Determinants of Health." *Lancet (London, England)* 379(9826): 1641–52. https://doi.org/10.1016/S0140-6736(12)60149-4.

Wang, Aolin, and Onyebuchi A. Arah. 2015. "G-Computation Demonstration in Causal Mediation Analysis." *European Journal of Epidemiology* 30(10): 1119–27. https://doi.org/10.1007/s10654-015-0100-z.

Westreich, Daniel, Jessie K. Edwards, Stephen R. Cole, Robert W. Platt, Sunni L. Mumford, and Enrique F. Schisterman. 2015. "Imputation Approaches for Potential Outcomes in Causal Inference." *International Journal of Epidemiology* 44(5): 1731–37. https://doi.org/10.1093/ije/dyv135.

Table 1 Effects of interest and their counterfactual definitions as risk differences

EFFECT	COUNTERFACTUAL DEFINITION	DESCRIPTION
Total effect (TE)	$E[Y_{xMx} - Y_{x*Mx*}]$	Overall effect of X on Y through all pathways
Pure direct effect (PDE)	$E[Y_{xMx*} - Y_{x*Mx*}]$	Effect of X on Y through pathways other than M
Pure indirect effect (PIE)	$E[Y_{xMx} - Y_{xMx*}]$	Effect of X on Y due to X affecting M and, subsequently, M affecting Y
Mediated interaction effect (MIE)	$E[Y_{xMx} - Y_{xMx*} - Y_{x*Mx} + Y_{x*Mx*}]$	Effect of X on Y due to both X causing M and X interacting with M
Proportion mediated (PM)	$\frac{E[PIE] + E[MIE]}{E[TE]}$	Proportion of the total effect that is due to an indirect pathway via M
Controlled direct effect (CDE)	$E[Y_{xm} - Y_{x*m}]$	Effect of X on Y when fixing M at a specific value for the whole population
Proportion eliminated (PE)	$\frac{E[TE] - E[CDE]}{E[TE]}$	Proportion of the total effect that was eliminated by fixing M

X = exposure, M = mediator, Y = outcome; x refers to the presence of health problems and x* to their absence; m is a fixed mediator value

Table 2 Frequency distribution, mean grade point average (GPA) and commonness of different educational pathways by background factors for the study population (N=73,072)

Covariate	Classification	%	Mean GPA	Completed education at age 23 (%) General	Vocational	None
Somatic condition at ages 10–16	No	59.9	7.67	52.3	36.9	10.7
	Yes	35.5	7.59	48.1	38.5	13.4
Mental disorder at ages 10–16	No	93.7	7.67	52.2	37.5	10.4
	Yes	6.3	7.18	27.9	39.2	32.9
Sex	Male	50.8	7.37	42.6	44.3	13.1
	Female	49.2	7.91	59.0	30.7	10.4
Urban-rural classification[a b]	Inner urban	19.3	7.80	59.9	26.2	14.0
	Outer urban	29.9	7.66	53.6	34.1	12.3
	Peri-urban	13.2	7.57	48.5	40.2	11.2
	Local centers in rural areas	6.1	7.59	47.7	41.9	10.4
	Rural areas close to urban	9.2	7.54	43.3	45.2	11.5
	Rural heartland	14.5	7.56	44.7	45.7	9.6
	Sparsely populated rural	7.8	7.56	42.0	47.4	10.7
Mother tongue	Finnish	93.5	7.64	50.5	37.7	11.8
	Swedish	4.8	7.73	55.9	36.7	7.4
	Other	1.7	7.26	42.5	33.5	24.0
Born in Finland	Yes	98.0	7.65	50.9	37.6	11.5
	No	2.0	7.25	39.9	36.4	23.7
Highest parental education[b]	Higher tertiary	15.3	8.37	84.4	11.4	4.2
	Lower tertiary	11.2	7.94	65.4	26.7	7.9
	Lowest tertiary	25.9	7.74	56.2	35.5	8.3
	Secondary	41.2	7.32	34.6	50.1	15.3
	Basic	6.4	7.02	24.5	47.5	28.0
Family type[b]	Two parents	81.0	7.70	53.3	36.9	9.8
	Mother only	16.6	7.37	39.6	40.2	20.2
	Father only	2.4	7.29	36.9	42.6	20.5
Birth year	6 categories (not shown)					
Region of residence	18 categories (not shown)					
Maternal age	Continuous[d]	28.6	0.14	29.4	28.0	27.3
Sibship size	Continuous	2.9	-0.07	2.8	3.1	3.1
Household disposable income[c]	Continuous	16,807	0.26	18,576	15,185	14,385
Total		100.0	7.64	50.6	37.6	11.8

[a]Based on a nationwide geographical classification system by the Finnish Environment Institute SYKE

[b]Mode at ages 10–14

[c]Average annual income at ages 10–14: annual measures were converted to 2012 euros and divided by the number of OECD (Organization for Economic Co-operation and Development) consumption units

[d]For continuous variables mean values, correlations with GPA and mean values by education at age 23 are shown

Table 3 Average effects (risk differences, RD, and their 95% confidence intervals, CI) of somatic conditions and mental disorders at ages 10–16 on the likelihood of completing no upper-secondary education by age 23, as mediated by grade point average (GPA)

	Somatic conditions			Mental disorders		
Effects	RD	Lower CI	Upper CI	RD	Lower CI	Upper CI
Adjusted effects (N=73,072)[a]						
Total effect	0.023	0.017	0.029	0.192	0.177	0.208
Pure direct effect	0.018	0.012	0.024	0.134	0.118	0.149
Pure indirect effect	0.005	0.004	0.006	0.037	0.034	0.041
Mediated interaction effect	0.000	0.000	0.001	0.022	0.015	0.029
Proportion mediated	0.224	0.161	0.305	0.306	0.267	0.350
Controlled direct effect A[b]	0.011	0.007	0.015	0.078	0.062	0.095
Proportion eliminated	0.532	0.327	0.706	0.592	0.507	0.673
Controlled direct effect B[c]	0.005	0.001	0.009	0.040	0.024	0.057
Proportion eliminated	0.773	0.550	0.943	0.791	0.700	0.875
Sibling fixed-effects (N=35,892)[d]						
Total effect	0.011	0.005	0.018	0.100	0.080	0.120
Pure direct effect	0.010	0.004	0.016	0.077	0.060	0.095
Pure indirect effect	0.001	0.000	0.002	0.009	0.007	0.012
Mediated interaction effect	0.000	0.000	0.001	0.013	0.009	0.018
Proportion mediated	0.096	-0.025	0.254	0.228	0.177	0.282

[a]Adjusted for sex, urban-rural classification, mother tongue, being foreign-born, parental education, family type, birth year, region of residence, maternal age, sibship size, and household disposable income

[b]GPA fixed at the 75th percentile of the GPA distribution (8.50)

[c]GPA fixed at the 90th percentile of the GPA distribution (9.09)

[d]Adjusted for sex

Table 4 Average effects (risk differences, RD, and their 95% confidence intervals, CI) of somatic conditions and mental disorders at ages 10–16 on the likelihood of completing the vocational track instead of the general track among persons who complete upper-secondary education by age 23, as mediated by grade point average (GPA)

	Somatic conditions			Mental disorders		
Effects	RD	Lower CI	Upper CI	RD	Lower CI	Upper CI
Adjusted effects (N=64,462)[a]						
Total effect	0.030	0.020	0.040	0.160	0.143	0.176
Pure direct effect	0.014	0.005	0.023	0.077	0.064	0.092
Pure indirect effect	0.016	0.010	0.022	0.090	0.079	0.100
Mediated interaction effect	0.000	-0.001	0.000	-0.008	-0.011	-0.005
Proportion mediated	0.529	0.363	0.742	0.515	0.459	0.576
Controlled direct effect A[b]	0.010	0.001	0.020	0.105	0.075	0.135
Proportion eliminated	0.649	0.324	0.952	0.341	0.193	0.505
Controlled direct effect B[c]	0.003	-0.003	0.011	0.059	0.034	0.084
Proportion eliminated	0.881	0.612	1.120	0.633	0.476	0.783
Sibling fixed-effects (N=31,921)[d]						
Total effect	0.009	0.001	0.018	0.065	0.042	0.087
Pure direct effect	0.007	-0.001	0.013	0.027	0.008	0.045
Pure indirect effect	0.003	-0.002	0.008	0.039	0.026	0.052
Mediated interaction effect	0.000	0.000	0.000	-0.001	-0.003	0.000
Proportion mediated	0.392	0.361	0.417	0.589	0.405	0.836

[a]Adjusted for sex, urban-rural classification, mother tongue, being foreign-born, parental education, family type, birth year, region of residence, maternal age, sibship size, and household disposable income

[b]GPA fixed at the 75th percentile of the GPA distribution (8.50)

[c]GPA fixed at the 90th percentile of the GPA distribution (9.09)

[d]Adjusted for sex

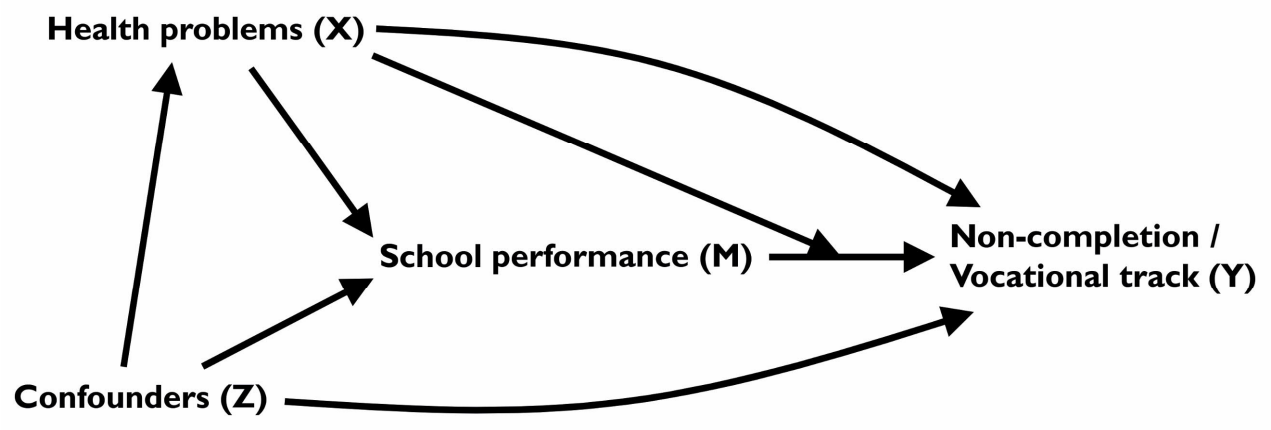
Table 5 Effects (risk differences, RD, and their 95% confidence intervals, CI) of alternative health problem exposures at ages 10–16 on non-completion of upper-secondary education by age 23 (n=73,072) and the choice of vocational track instead of general track (n=64,462), as mediated by grade point average. Prevalence of the health problem indicator in parentheses

		Non-completion			Vocational track		
Health problems	Effects[a]	RD	Lower CI	Upper CI	RD	Lower CI	Upper CI
Somatic treatment during 4 or more years at ages 10–16 (7.3%)	Total effect	0.027	0.016	0.039	0.033	0.013	0.053
	Pure direct effect	0.022	0.011	0.033	0.013	-0.003	0.031
	Pure indirect effect	0.005	0.003	0.007	0.021	0.010	0.032
	Mediated interaction	0.000	0.000	0.001	-0.001	-0.003	0.000
	Proportion mediated	0.198	0.099	0.331	0.691	0.321	1.208
Complex chronic condition (2.9%)	Total effect	0.030	0.013	0.047	0.025	-0.004	0.055
	Pure direct effect	0.030	0.012	0.047	0.025	0.000	0.050
	Pure indirect effect	0.000	-0.003	0.004	0.000	-0.016	0.018
	Mediated interaction	0.000	-0.001	0.001	0.000	-0.001	0.001
	Proportion mediated	-			-		
Internalizing: depression or anxiety (2.5%)[b]	Total effect	0.195	0.175	0.215	0.137	0.111	0.165
	Pure direct effect	0.151	0.133	0.169	0.077	0.054	0.098
	Pure indirect effect	0.033	0.028	0.039	0.069	0.052	0.086
	Mediated interaction	0.011	0.005	0.016	-0.009	-0.015	-0.005
	Proportion mediated	0.225	0.187	0.265	0.439	0.344	0.551
Externalizing: attention deficit hyperactivity disorder or conduct disorder (1.5%)[b]	Total effect	0.289	0.261	0.319	0.315	0.281	0.350
	Pure direct effect	0.196	0.163	0.233	0.167	0.121	0.213
	Pure indirect effect	0.076	0.068	0.084	0.185	0.164	0.206
	Mediated interaction	0.016	-0.005	0.035	-0.037	-0.062	-0.016
	Proportion mediated	0.321	0.237	0.394	0.472	0.369	0.586

[a]Adjusted for sex, urban-rural classification, mother tongue, being foreign-born, highest parental education, family type, birth year, region of residence, maternal age, sibship size, and household disposable income

[b]The effects of internalizing and externalizing were mutually adjusted for each other

Figure 1 Directed acyclic graph (DAG) depicting the causal assumptions. X, M and Y represent the exposure, mediator and outcome, respectively. Health problems are assumed to moderate the association between school performance and non-completion/vocational track (intersecting arrow). The associations are assumed unconfounded conditional on Z, which represents a set of exposure–outcome, exposure–mediator and mediator–outcome confounders (see Methods section for details)



Supplementary Table 1 Prevalence (%) of different types of health conditions within the broad groups of somatic conditions and mental disorders

Health condition group	ICD-10 code	Prevalence
Somatic conditions		35.5
Infectious and parasitic diseases	A00-B99	2.6
Neoplasms	C00-D48	1.5
Diseases of the blood and blood-forming organs; immune system disorders	D50-D89	0.5
Endocrine, nutritional and metabolic diseases	E00-E90	3.1
Diseases of the nervous system	G00-G99	2.8
Diseases of the eye and adnexa; diseases of the ear and mastoid process	H00-H95	6.0
Diseases of the circulatory system	I00-I99	0.9
Diseases of the respiratory system	J00-J99	11.9
Diseases of the digestive system	K00-K93	5.1
Diseases of the skin and subcutaneous tissue	L00-L99	3.7
Diseases of the musculoskeletal system and connective tissue	M00-M99	6.1
Diseases of the genitourinary system	N00-N99	3.0
Congenital malformations, deformations and chromosomal abnormalities	Q00-Q99	2.6
Mental disorders		6.3
Mental and behavioral disorders due to psychoactive substance use	F10-F19	0.5
Schizophrenia, schizotypal and delusional disorders	F20-F29	0.2
Mood disorders	F30-F39	1.9
Neurotic, stress-related and somatoform disorders	F40-F48	1.8
Behavioral syndromes associated with physiological disturbances and physical factors	F50-F59	0.7
Disorders of adult personality and behavior	F60-F69	0.1
Disorders with onset usually occurring in childhood and adolescence; unspecified	F90-F99	2.9
Internalizing		2.5
Depression	F32-F33[a]	1.9
Anxiety	F40-F42	0.9
Externalizing		1.5
Attention deficit and hyperactivity disorder / hyperkinetic disorder	F90	0.5
Conduct disorder	F91-F92	1.1

[a]Excluding codes F32.3 and F33.3 (depression with psychotic features)

Supplementary Table 2 Average effects (risk differences, RD, and their 95% confidence intervals, CI) of somatic conditions and mental disorders at ages 10–16 on the likelihood of listing the vocational track (instead of the general track) as the first preference when participating for the first time in the upper-secondary school application process, as mediated by grade point average; persons born in 1992-1993 (n=24,175)

	Somatic conditions			Mental disorders		
Effects[a]	RD	Lower CI	Upper CI	RD	Lower CI	Upper CI
Total effect	0.026	0.013	0.038	0.154	0.132	0.175
Pure direct effect	0.011	0.001	0.020	0.065	0.045	0.084
Pure indirect effect	0.014	0.007	0.022	0.098	0.084	0.112
Mediated interaction effect	0.000	-0.001	0.000	-0.009	-0.015	-0.004
Proportion mediated	0.573	0.323	0.094	0.578	0.492	0.672

[a]Adjusted for sex, urban-rural classification, mother tongue, being foreign-born, highest parental education, family type, birth year, region of residence, maternal age, sibship size, and household disposable income

Supplementary Table 3 Average effects (risk differences, RD, and their 95% confidence intervals, CI) of somatic conditions and mental disorders at ages 10–16 on the likelihood of completing no upper-secondary education by age 23, as mediated by failure to enter the preferred study program in the upper-secondary school application process; persons born in 1992-1993 (n=24,175)

	Somatic conditions			Mental disorders		
Effects[a]	RD	Lower CI	Upper CI	RD	Lower CI	Upper CI
Total effect	0.019	0.012	0.027	0.166	0.146	0.186
Pure direct effect	0.018	0.010	0.025	0.150	0.130	0.170
Pure indirect effect	0.002	0.000	0.003	0.011	0.009	0.014
Mediated interaction effect	0.000	0.000	0.000	0.005	0.000	0.010
Proportion mediated	0.092	0.008	0.187	0.096	0.063	0.136

[a]Adjusted for sex, urban-rural classification, mother tongue, being foreign born, highest parental education, family type, birth year, region of residence, maternal age, sibship size, and household disposable income