A critical review of the literature on the relationship between school quality and health inequalities

Anthony M. Garcy1,* and David C. Berliner2

1CHESS, Center for Health Equity Studies Stockholm University and Karolinska Institute, Stockholm, Sweden, 2Arizona State University, Tempe, USA

Robust evidence suggesting a strong association between greater educational attainment, better health and lower mortality, has led to speculation that the quality of schooling can also have effects on health. This review critically summarises findings from 15 studies in a growing area of research concerning the effects of school quality on health. Findings suggested positive, long-term benefits of high-quality pre-school. Other findings suggested that higher teacher wages, lower pupil–teacher ratios, a longer school year, and higher college selectivity had mostly positive long-term effects on health and mortality. Several studies found that school quality modified the effect of years of completed education on various health outcomes. Some measures of school quality including smaller class size in relation to mortality, and higher college selectivity in the case of smoking were not consistently related to better health. While studies varied in their consistency and significance, the weight of the evidence together, suggests that some health inequalities over the life course were explained partly by differences in school quality. This may be related to improved cognition, occupational characteristics, and the incomes of those exposed to better quality schooling. Direct health knowledge and behaviour may also play a role.

Introduction

This review considers a well-established belief across many societies, namely, that a person’s health and educational achievements and attainment are correlated. It is not difficult to imagine that such a relationship might even be causal, though only correlational data supports this belief. Poor health is quite likely to detract from school achievement, as in the case of recurring otitis media that hampers hearing, or chronic asthma that weakens children and contributes to school absence. Good health in contrast, is likely to promote higher achievement in its most obvious form by high rates of student attendance and sufficient energy for learning.

This correlational relationship is also likely to be bi-directional. If this is true, then school policies, programmes, and quality influence health outcomes. In this critical review, instead of looking at how health affects schooling, we look at how different characteristics of educational systems can affect health. If that relationship can be demonstrated, it provides another reason for why high-quality public schooling for all should be a national goal.

*Corresponding author. CHESS, Center for Health Equity Studies, Stockholm University and Karolinska Institute, Sveaplan, Sveavagen 160, Floor 5, 10654 Stockholm, Sweden. Email: anthony.garcy@chess.su.se

© 2018 British Educational Research Association
Robust, consistent evidence has suggested that more advantaged groups enjoy better health (Ross & Mirowsky, 1999), and longevity (Lleras-Muney, 2005; Grossman, 2006) than less advantaged groups. Findings from international studies have demonstrated the existence of a graded association between greater individual-level educational attainment and better health (Lahelma et al., 2004; Avendano et al., 2009). In Japan (Kagamimori et al., 2009) and in the USA this relationship is more complex as it has been shown to be nonlinear (Backlund et al., 1999; Rogers et al., 2010; Zajacova et al., 2012; Zajacova, 2012), and dependent on gender and race/ethnicity (Kimbro et al., 2008; Zajacova & Hummer, 2009; Ross et al., 2012). The association is also likely to be more consistent in southern and western European countries than in the Nordic welfare states where lower levels of education may not confer the same level of health disadvantage (Avendano et al., 2009).

The Black report (Townsend & Davidson, 1982) cogently summarised four classes of theory that offered divergent explanations for health differences between individuals and social groups. They included statistical artifact, cultural/behavioural theories, selection related explanations, and structural/materialist theories. Since this time, little evidence has been found to support the theory that health differences are predominantly attributable to artifact. Similarly, cultural/behavioural theories that are strongly promoted in the health-care sector have been criticised for their over reliance on the primacy of individual-level agency in health related behaviour. Typically, they make no contextual link between behaviour and social structural influences or constraints.

In relation to educational processes, selection and structural theories are the most germane. Although there have been refinements to these two opposing theories over time, basic versions of each suggest that differences in health can either be explained by individual differences in genetic endowments (e.g. intelligence) and the initial distribution of physical health, or by differential access to resources (e.g. high-quality education) that is determined by higher/lower position in social structures.

Gottfredson (2004) and others (Gottfredson & Deary, 2004) have argued strongly that general intelligence is the fundamental cause that explains differences in health. Numerous studies have shown that high IQ individuals are healthier over the life course and they live longer than individuals with lower IQs. This positive association has been attributed frequently to healthier behaviour including better compliance with medical routines, or better decision making in health related matters (Gottfredson, 2004). Most studies have inferred this to be true but have not explicitly demonstrated it. Some have also suggested, based mostly on findings from twin studies that assume identical family backgrounds that general intelligence is mostly heritable and is not affected permanently by shared rearing and environmental conditions (Gottfredson, 2004). However, this may not apply to populations with very low social position. Turkheimer et al. (2003) have shown large effects of shared family background on cognition in impoverished families.

The twin study findings led to the development of strong convictions among many that biometrically informed research would eventually uncover the genotypes behind general intelligence and other complex human behaviours (Turkheimer, 2000). Studies that have utilised genome wide association databases have
reported statistically significant associations between specific alleles and characteristics of individuals (e.g. educational attainment) (Rietveld et al., 2013; Belsky et al., 2016) and some indicators of cognitive ability (Belsky et al., 2016). To date, only a small number of alleles have been identified. Even when these effects have been considered together, the overall effect size has been very small (e.g. educational attainment coefficient of determination \( \approx 2\% \)) (Rietveld et al., 2013).

While it is clear that genes are related to behaviour in a very general sense, Nisbett et al. (2012), Turkheimer (2016) and others have argued that biometrically informed studies have not produced a revolution of understanding concerning genetic links to complex human behaviours or intelligence. Turkheimer (2016) summarised the increasingly accepted position among many conducting biometrically informed research. He wrote:

> Note that Loehlin and Martin’s alternative to genetic was not ‘environmental’ but ‘phenotypic.’ The nature/nurture debate was supposed to be about whether differences in behavior were better characterised as genetic or environmental, but that turned out to be the wrong question. Everything is both genetic and environmental. The deeper question involves determining what is to be learned by examining phenotypic phenomena at a genetic level of analysis.

As the debate about intelligence evolves from a genetically deterministic reductionism to one about phenotype shaped by relative environmental context, it can be acknowledged that more intelligent people are likely to enjoy better health than less intelligent people. Yet this assertion is not contradicted by the statement that social structural processes shape both intelligence and health among groups of individuals who have been exposed to different environments (e.g. different family backgrounds, schools and neighbourhoods).

Researchers, who have given more weight to social structural explanations for persistent group health differences in highly developed countries, have speculated that educational policy and characteristics of highly developed countries’ educational systems could have effects on population and individual-level health (Low et al., 2005; Muennig & Woolf, 2007; Fiscella & Kitzman, 2009; Ruglis & Freudenberg, 2010; Cohen & Syme, 2013) irrespective of individual-level cognitive ability. In the USA (Olshansky et al., 2012), and in Europe, some have suggested that growth over time in life expectancy and health differentials between the highly educated and the least educated could be reduced by implementing various ‘educational enhancements’ (Low et al., 2005). This includes proposals to increase the mandatory school leaving age in some countries (e.g. Sweden), lower the starting school age (e.g. the Netherlands), offer of high quality day care to all parents, and guaranteed community college attendance to those who meet achievement standards (e.g. USA).

This reasoning emerges from the idea that structural characteristics of a country’s education system like those related to school quality (e.g. class sizes, teacher quality and the presence of early tracking) influence the educational outcomes of attending students independently of their background characteristics. Evidence from cross-national, comparative work using international assessment data has supported this.
claim in the context of early tracking when high achievement inequality within countries occurs (Alegre & Ferrer, 2010; Van de Werfhorst & Mijs, 2010; Condron, 2011). However, most if not all of this work is based on cross-sectional data. Hence, causal inference remains problematic.

The idea that changes in education policy could produce changes in aggregate-level health was likely first proposed in the early 1980s in the context of public health in developing countries (Grosse, 1982). Much of the focus in the scholarly literature has been on the attained quantity of education, and subsequent health effects. These studies examined relative differences in cohort mortality rates (Lleras-Muney, 2005; Mazumder, 2008; Clark & Royer, 2010; Meghir et al., 2012), general health (Mazumder, 2008; Silles, 2009; Clark & Royer, 2010), and self-reported health (Arendt, 2005; Silles, 2009; Clark & Royer, 2010) in the USA and in a small number of European countries. Largely because of a scarcity of data about education system configurations/policies that provide both measures of school quality and health indicators for individuals, only a small number of empirical studies exist.

Broadly conceived, school quality refers to the structural and organisational features of educational systems (e.g. the effect of tracking, the degree of segregation in schools, the availability of early high-quality childhood/pre-school education, and college selectivity). Quality is also determined by the resources available to schools including economic and financial support (affecting class size, per-pupil instructional expenditures, teacher pay, school-year length, etc.) and human resources (e.g. teacher quality, principal competence). Research on the effects of school quality on health has largely been conducted in four areas including: early childhood education; school resource effects; school racial and ethnic composition; and college selectivity. This review critically assesses the evidence relating various aspects of school quality to health differences. The evidence suggests that school quality can explain some health disparities/inequalities between more and less advantaged groups.

Method

A bibliographic search was undertaken to identify all studies that have been completed in the areas of demography, economics, education, epidemiology, psychology, public health and sociology between 1970 and the end of 2014. Searches were conducted in the ERIC, PUBMED, Web of Science and Worldcat search engine databases. A two-stage search process was used. Single education related keywords were first entered (e.g. education quality, school quality, education policy, education systems, early education, pre-school, small class size, class sizes, health, school racial composition and college selectivity). The second stage employed a Boolean search strategy where educational keywords were followed or preceded by health relevant keywords (e.g. student health, health, health outcomes, health effects, public health effects, mortality and disability). We selected empirical articles based on quantitative cross-sectional or longitudinal data that had been published in a peer review journal. Heavy weight was given to studies that focused on the structural aspects of school quality mentioned earlier. Those
that assessed the effects of school quality measured early in the life course on later life course health outcomes were considered the most relevant. Fifteen studies were selected and reviewed from a group of 25 unique studies. The search results and screening decisions are shown in Figure 1.

This review excluded studies that assessed the efficacy of health education or health interventions implemented directly in schools, e.g. in-school health programmes (Symons et al., 1997), nutrition programmes (Figlio & Winicki, 2005; Hinrichs, 2010), and programmes that may have affected smoking/drug/alcohol use or physical activity. This literature has not usually assessed the long-term health or educational effects of these interventions over the life course. Moreover, these studies have almost exclusively focused on the narrow period between pre-adolescence and adolescence. In addition, several comprehensive reviews have already summarised this literature (Fletcher et al., 2008; Bonell et al., 2013a,b). A review based on qualitative studies of substance use, diet and sexual health was not included (Jamal et al., 2013). We also excluded articles that had no measure of school quality (Apps et al., 2013), had used an aggregate-level

Figure 1. Study selection flow chart
international student achievement test score as a measure of school quality (Jami-
son et al., 2007), or included no health outcome (Gomes-Neto et al., 1997).

**High-quality pre-school and health**

We found two small random controlled trials (RCTs) that assessed the effects of early, high-quality pre-school programmes on the later health of the participants (Muennig et al., 2009; Muennig et al., 2011a). Findings from the High Scope Perry Preschool Program \( (n = 123) \) where all teachers were highly qualified (all had both a master’s degree and child development training) suggested some limited health benefits. The study randomly assigned 58 children into the treatment group and 65 children to the control group. Children were followed from the ages of 3 or 4. One-hundred and twelve survived or could be located at the age of 40 when they were surveyed. Information was obtained from the participants on general health status, specific health conditions, health behaviours including tertiary health-care use, preventative medical care use and health risk behaviours.

At age 40, those who were exposed to the programme showed a few statistically significant differences in health behaviours in contrast to the control subjects. There was more seatbelt use at age 27 (23% more), less use of alcohol (12% less), and no heroin use among the exposed subjects. The relatively young age of the participants made it difficult to detect health differences between the groups, which could emerge later. However, earlier related findings suggested that individuals in the treatment group had higher regular high school graduation rates (65% vs 45%), were more likely to have employer sponsored health insurance (65% vs 45%), and were more likely to have earned $20K in income by age 40 (60% vs 40%) (Schweinhart et al., 2005). Importantly, the study also found that some of the differences in health behaviour were explained by the educational attainment and the availability of health insurance among the treated group.

A second small \( (n = 104) \), more intensive, high-quality RCT of a pre-school intervention called the Carolina Abecedarian Project suggested improved health behaviour among exposed subjects, as well as lower rates of pregnancy and depression at age 21 (Muennig et al., 2011a). However, low statistical power made it difficult to detect significant programme effects on later health behaviours and measures of health outcomes for the exposed individuals. Other benefits reported in an earlier study included better cognitive performance and greater college enrollment, which were not explained by infant IQ, scholastic performance at age 15, or access to health care (Campbell et al., 2002; McLaughlin et al., 2007).

In these studies, infant participants were all healthy but had been identified as high risk for poor educational outcomes. The intervention programme used a specially designed curriculum to stimulate language and cognitive development from infancy until the age of 8. Programme participants received pediatric care on site. Parents of the treatment group children also received support and social services, as needed. Control group subjects received medical care in a local, low-cost well-baby centre. Both groups received nutritional supplements and social work services to ensure that any observed program benefits resulted from the curriculum rather than differences in early childhood nutrition.
We found three studies that searched for effects from smaller class size on health outcomes. Two longitudinal studies using similar data examined mortality (Muennig et al., 2011b) and disability status and disability claims in the USA (Wilde et al., 2011). A third cross-sectional study examined class size and mental health in Sweden (Jakobsson et al., 2013).

Data from the multi-school, RCT STAR (Student Teacher Achievement Ratio) project were used to assess the effect of class size reductions on mortality to age 29 (Muennig et al., 2011b). At the beginning of each school year, Tennessee state teachers and students from kindergarten to grade 3 were randomised within schools and assigned to classrooms of 13–17 and 22–25 students respectively. In some of the larger classes an aide was also randomly assigned. School selection was not, however, random and was instead based upon the size and the willingness of the school to participate. The experiment was conducted in 328 classrooms in 79 different school districts located in urban, suburban, and rural areas. These data were then linked via a probability match to the US National Death Index between 1985 until the end of 2007 so that any deaths among the STAR participants could be logged. The final sample included 11,240 participants with a total of 141 deaths.

The findings from the main Cox regression analysis suggested that children assigned to the smaller classes had higher mortality (Hazard Ratio (H.R.) 1.58; 95% Confidence Interval (C.I.) 1.07–2.32) in contrast to those assigned to the larger classes. The presence of an aide made no difference. In subgroup analyses this effect was mostly driven by males (H.R. 1.73; 95% C.I. 1.05–2.85), white and Asian students (H.R. 1.68; 95% C.I. 1.04–2.72), and those who were not likely to have come from a poor family (H.R. 2.20; 95% C.I. 1.06–4.57). Also of interest was the finding that better school performance measured by test performance after the fifth grade was associated with smaller class size and lower mortality. However, mediation analyses run on a subsample of about half the students where test scores were available suggested that better school performance after grade 5 did not reduce the higher mortality effect associated with smaller class size. In fact, it increased marginally.

A battery of additional analysis was done to assess the robustness of the main Cox regression results. Other estimation methods including ordinary least squares (OLS) and probit were used. Both the Cox models and the additional estimations adjusted the standard errors for classroom clustering and a fixed school-grade entry effect. Tests that confirmed that the proportionality assumption could not be rejected in the case of the Cox analysis were conducted as well. Because of the small number of deaths, additional sensitivity analyses were undertaken. These simulated how the results would be affected if the number of deaths in each of the treatment groups changed to a marginal extent. These supplementary analyses did not suggest alternative findings when compared with the main analyses.

Given previous well-documented findings that have shown that students assigned to the smaller classrooms in project STAR experienced a number of positive life course outcomes, the authors speculated that the counterintuitive findings concerning mortality might be attributable to non-cognitive behavioural changes among students in the groups most at risk (e.g. more experimentation with drugs/drinking). In
fact, deaths attributable to poisonings, alcohol, drugs, intoxicated driving, and firearms accounted for most of the deaths in the sample. The unadjusted mortality figures also gave some support to this claim. It was further speculated that these risk behaviours outweighed the protective effects of the enhanced cognitive returns the class size reduction had on the achievement and schooling of the 42 students who died.

These results should be considered carefully for several reasons, some of which the authors acknowledged. Some switching between the treatment groups occurred at later points in each school (22.5%). Because of the switching an intent-to-treat analysis was done such that some partially treated children who switched were counted as untreated and vice versa. This led to compositional differences between the groups over time. Therefore, some analyses included controls for some student characteristics including race/ethnicity, free lunch status, and whether the student was in an urban, suburban, or rural school, but it is not completely evident if these adjusted model results were presented in the paper.

Although random assignment of teachers and students within schools should have overridden concerns about potential bias owing to student and teacher characteristics, the fact remains that the school sample was not random. This meant that small schools, those that had resource constraints, or those that disapproved, did not participate (Biddle & Berliner, 2002). Comparisons made between the participating schools and the rest of the Tennessee state public schools revealed that participating schools were larger and had lower mathematics and reading achievement scores (Finn et al., 2007). This limits the external validity of the study’s findings. In addition one needs to consider that many of the treated classes never showed the effects of treatment (Konstantopoulos, 2011). As Konstantopoulos (p. 31) said:

Although students benefit considerably from being in small classes in many schools, in other schools being in small classes is either not beneficial or is a disadvantage. Small class effects were inconsistent and varied significantly across schools in all grades indicating a small class by school interaction.

Finally, it was pointed out that reduced mortality effects could emerge later given the small number of deaths and the relatively young age of the participants during the follow-up. With these findings and critiques in mind, we examine a second study based on the project STAR data.

In this study Wilde et al. (2011) examined the effect of small class size on the reporting of disability claims shortly after secondary school completion. Their findings suggested no statistically significant differences in mean earnings, mean years of employment, or mean disability claims between the small class size group, the regular class size group, and the regular class size group with an aide. The project STAR data were probabilistically matched to US Social Security Administration (SSA) records that contained annual individual-level information on mean and annual earnings, years of employment, annual employment and disability claim status during the 1997–2008 period. Only 54% of the 11,601 STAR participants were matched to these SSA records. The STAR participants generally graduated from secondary school toward the beginning of this period. Some of the younger participants may have even graduated from college.
Because of switching between treatment groups that occurred in the study an intent-to-treat regression analysis with robust standard errors to correct for school and classroom clustering was run. The study, however, suffers from several problems, some of which are acknowledged by the authors. For example, the match is likely to have produced a subsample of participants that was not representative of the entire STAR sample. Secondly, the low match rate hindered statistical power to detect treatment effects. Finally, given the repeated measures they had (12), it is unclear why they did not use an analysis designed to examine trajectories of earnings, employment and disability status according to treatment group membership. Potentially, this analysis might have been more informative in terms of understanding the overall trends present in these data.

An additional study of class size and health related outcomes exists (Jakobsson et al., 2013). Using cross-sectional Swedish survey data, collected in 2008 from ninth graders \( n = 2755 \), in the county of Värmland, these investigators showed no effect of class size on various measures of mental health. A variety of modelling strategies were used that corrected for clustering at the school and classroom levels including probit, school fixed effects, an instrumental variable approach and a regression discontinuity design. There were very few substantive differences between each of the sets of analyses, net of the control variables. As the study’s authors acknowledged, the findings must be contextualised given consistent findings concerning non-health related benefits of smaller class size. It is also unclear how well these findings can be generalised to other municipalities. In Värmland, a higher amount of funding per student was allocated based on whether the student received a disadvantaged classification. In other municipalities how students were classified as disadvantaged varied.

School resources—pupil–teacher ratios, school-year length, and teacher wages

Sansani (2011) used a quasi-experimental approach, the Biennial Survey of Education data, and American Census data, to construct synthetic birth cohorts for the estimation of death rates based on the approach used by Lleras-Muney (2005). Sansani demonstrated that state-level indicators of school quality had a positive relationship to health. Death rates decreased by about 1.5 percentage points when individual measures of school quality (pupil–teacher ratios, and the length of the school term) increased by 1 standard deviation. It was also shown that higher relative teacher wages were associated with a lower state mortality–schooling gradient net of the other school quality indicators and average state income. Deaths decreased by 1.9 per 1000 people for each additional year of education as relative wages of teachers increased by 1 standard deviation.

Sansani suggested that the study’s findings must be considered cautiously. In many states improvements in school quality were not likely to be independent of concurrent investments in childhood health. Effort was made, however, to minimise this concern with the inclusion of state of birth and census pair fixed effects, but, another problem with this study was that data on family background and ability were not available. Nevertheless, all three indicators of school quality were significant predictors of the state mortality–schooling gradient.
Similarly, Frisvold and Golberstein (2011) used a quasi-experimental design and data from the representative (US) National Health Interview Surveys. The study found modest, augmenting effects of school quality for a broader variety of measures of health in later life (1984–2007). These included better self-rated health, a lower probability of obesity and smoking, and slightly reduced mortality. The study included Black Americans \((n = 42,902)\) who attended highly segregated primary and secondary schools between 1936 and 1966 in 18 southern states. Because black families and communities had negligible, if any, influence on public school financing at this time, a focus on this group of students was likely to better identify the magnitude of the effect of school quality on health. Additionally, eradication of several diseases that could have influenced school attendance including hookworm, malaria, yellow fever and pellagra had occurred prior to the birth of those in the oldest cohort, i.e. the 1930 cohort. Thus, the findings are unlikely to be confounded by major public health improvements. Measures of school quality included state and year specific reductions in the pupil–teacher ratio, increases in the length of the school year and average state teacher’s wages.

The authors tested the specific hypothesis that the health returns to a year of schooling were greater from a higher quality school than from a lower quality school. They reasoned that this specification, i.e. the interaction between attained education level and the state school quality indicators would be less biased by omitted variables. Importantly, they stated that reduced form models that included only the main effects of the school quality variables had no statistically significant relationship with the health indicators. Thus they speculated that school quality modified the effect of education by improving cognitive ability.

The study provided health related estimates of the return to schooling, modified by the effect of an increase in school quality. For example, between 1930 and 1950 the sample average pupil–teacher ratio decreased by 10 units. At this mean pupil–teacher ratio decrease, an additional 4 years of schooling was estimated to improve self-rated health by 0.38 points. It was projected that if a similar improvement (reduction) in the pupil–teacher ratio occurred, an additional 4 years of schooling would raise self-rated health by 0.46 points. Similarly, the probability of being obese decreased more than a percentage point for an additional 4 years of education when accompanied by a 10 unit decrease in the pupil–teacher ratio. Clearly, the modifying effects of an increase in education quality with additional schooling were meaningful to health but modest in size.

Another set of analyses included the same health outcomes and the mediating effects of occupation, income, insurance status and marital status to assess how higher quality schooling might produce better health. Findings indicated that a part of the better self-rated health relationship was attributable to income, to a lesser degree to occupation and to access to health insurance. For the school quality/obesity relationship, occupation explained part of the association while the other mediators had no role. The school quality/mortality relationship was mediated by occupation and income only. However, it is not evident how the magnitude of these estimates can be interpreted. Interestingly, the social network measure (marital status) had no mediating effect in any of the models. Based on these results it was speculated that the unmediated portion of the school quality/health relationship was likely a result of
the cognitive advantage that higher quality schooling imparted. However, they had no direct evidence of this.

The authors pointed out that the results could have been sensitive to selective migration, where families that were concerned with education quality, or greater access to health care, might have moved out of certain states. In fact, mass migration of blacks to northern states did occur during the coverage period of this study. Likewise, the mortality estimates could have been sensitive to higher mortality among those individuals with lower quality education.

School segregation and health

The literature to date has often been contradictory concerning predictions about the expected health effects of school segregation. It has been suggested that mental health should be poor among students in low-socioeconomic status (SES) schools where there are high concentrations of minority students and higher rates of violence and disorder. Such conditions can be stressful and could potentially contribute to higher levels of depression (Mazza & Reynolds, 1999) and somatic health problems among students (Natvig et al., 1999).

In segregated, but high-SES schools, it is less clear how the health of students should fare. Minority students could benefit from access to better educational resources and opportunities that can positively impact educational achievement (Crosnoe, 2009) and later adult health. At the same time, it has been suggested that minority students attending predominantly majority group schools have a greater risk of feeling marginalised, are less attached to school and more likely to experience stress. This could lead to concurrent mental and somatic health problems (Crosnoe, 2009). Strong peer bonds between minority group students attending segregated, high-SES schools may produce beneficial health effects because minority group students are less likely to adopt the smoking and drinking behaviours of majority group students (Walsemann & Bell, 2010).

The theoretical literature has clearly produced competing predictions about the health effects of segregation in different school settings. As a reflection of this debate, an emergent literature of empirical studies done in the USA based on both cross-sectional and longitudinal data offers a mixed, inconsistent picture about the effects of school and curriculum-level segregation in relation to adolescent and adult health. Disparity in the findings among these studies is partially attributable to differences between transient and longer-term health outcomes among study subjects.

A cross-sectional study using one wave of data from the US National Longitudinal Study of Adolescent Health (AddHealth) tested two competing hypotheses related to high segregation within English curriculum tracks (Walsemann & Bell, 2010). The authors suggested high segregation would either: (1) lead to more smoking and drinking among black students relative to majority students, or (2) lead to less smoking and drinking as black students banded together and became less likely to adopt the health damaging behaviours of the majority group students. Findings suggested that high levels of school segregation in the English curriculum had mixed effects on the probability of smoking and drinking that depended on race/ethnicity and gender. White girls relative to black girls in highly segregated schools had a higher probability
of drinking and smoking (0.54) while black girls in highly segregated tracks appeared to benefit healthwise as they had a lower probability of this same behaviour (0.25). There were no similar findings for male students.

AddHealth is a nationally representative sample of adolescent students in 80 schools, in grades 7–12, during the 1994/95 school year. Survey data were collected from students, parents and school administrators. It also included retrospective transcript information. The final analytical sample included information from 6889 students, from a little over half of the available sample schools \((n = 47)\). It was not evident why some schools were excluded nor was it clear how many students were excluded from the original sample.

An index of dissimilarity was calculated based on the proportions of black and white students who were enrolled in different tracks of a school’s English curriculum. The authors reasoned that this would better reflect the effect of the school’s racial/ethnic composition because most students were required to take English as a subject. A dose response relationship was found between this index and the probability of smoking and drinking. Three dummy variables were created for low, medium and high values of the index. Interactions between the different levels of curriculum segregation and black/white racial status were also created to assess whether the relationship between segregation and the health behaviour outcomes varied by racial status. Model control covariates included the student-level characteristics of gender, age, race, school aspirations, family structure and family SES. A set of school-level characteristics were measured including school-level SES, the mean percentage of white students in the school, school size, school regional location, school type (private/public) and school urbanicity. Two-level general linear models that included a random intercept were used to address the matter of clustering of students within schools. Analyses were stratified by gender.

Given the cross-sectional design of the study and the non-specific direction of the index of dissimilarity, it was not possible to determine from the study if highly segregated referred to a high proportion of white or black students in a given set of tracks. Hence, it was not clear if the higher probability of drinking and smoking for white females in schools with highly segregated curriculum tracks was the result of a situation where the students were found among mostly white or mostly black peers. Similarly, it was equally difficult to assess whether the apparent health benefit of this same exposure for black girls was the result of being embedded mostly in tracks with black or white peers. Thus, it is difficult to interpret the findings in relation to the original hypotheses. The study also failed to give any explanation for why no similar effect was found for white or black males. Finally, the loss of almost half of the school sample severely limited the validity of any generalisations that could be made from the study.

School racial composition has also been linked to student mental health. Net findings from a cross-sectional study using wave I (1994/95) of AddHealth suggested that the level of depressive symptoms experienced by black students increased as the percentage of white students in a school increased (Walsemann et al., 2011). Similarly, the level of depressive symptoms among white students was higher in schools with a lower percentage of white students. The difference between white and black students in the predicted level of depressive symptom reporting was generally statistically different until controls for perceived discrimination and school attachment were
included in the statistical models. This suggested that poor mental health among black students was explained by a sense of discrimination and marginalisation. It was not clear if this was also the case among white students in schools where they were a minority. The authors reported a similar effect for somatic symptoms. However, plots showed that relative to white students, symptoms among black students were in fact lower at all levels of segregation. But it was evident that somatic symptom reporting increased with higher levels of segregation. There were no similar health findings for students of other racial/ethnic backgrounds. Small subgroup sample sizes may have hindered the study's ability to detect these effects.

The final analytical sample included data from 18,419 students in 132 junior/senior schools. Depressive symptoms were measured via a 19-item scale that was summed and transformed to address a slightly skewed distribution. Somatic symptoms were measured on a 12-item scale and dichotomised into a high/low indicator. School racial composition included the percentage of non-Hispanic, white students. Self-identified race/ethnicity was aggregated to the school level and weighted so that the measure was representative. Student-level control covariates included age, race/ethnicity, immigrant status, family structure and SES, perceived discrimination, and the degree of school attachment. School SES, urbanicity, region location and school type were measured at the school level.

A two-level linear mixed modelling strategy was used to assess the effect of school segregation on depressive and somatic symptoms. Sensitivity analyses were also run to assess the influence of schools that had either 0% or 100% of white students. Stratified analyses by racial/and ethnic group independently assessed interaction effects between the segregation measure and each racial/ethnic grouping. The authors reported that there were no discrepant findings in these additional analyses.

Several limitations of the study were noted including the possibility that the higher depressive symptoms among the black students who attended predominantly white schools might not translate into negative health effects later in the life course. In fact, it was suggested that black students might derive other health and educationally related benefits from attendance at a highly segregated school with a large proportion of white students even if there were short term, negative health consequences. The analyses also failed to employ post-hoc or planned contrasts at each level of segregation. Therefore, it was impossible to know if all plotted differences in each health outcome between black and white students were statistically different from one another. It is also possible that compositional thresholds existed such that some combinations were problematic for health while others were not. Finally, the authors suggested that heterogeneity in the composition of the other racial/ethnic groups might have masked potential health differences among the other groups.

A longitudinal study that used all four waves of the AddHealth data set (Walsemann et al., 2011) worked with four measures of depressive symptoms. These were obtained as early as age 14 and as late as age 34. Hispanic and Asian students initially reported higher depressive symptoms if they had attended a school in 1994/95 with a low percentage of whites but a high percentage of minority students. Over time symptom reporting declined for both groups and may have even been lower for Hispanics relative to white and black students in similar schools with a low percentage of white
students. School racial composition did not adversely impact the mental health of these students over time.

For black students who attended a predominantly minority school, depressive symptom reporting was initially lower and then became about the same relative to white students who had attended schools with a similar racial composition. There was no formal statistical test of the differences over time. However, a comparison with black students who had attended schools with a large proportion of white students appeared to show that the level of symptom reporting was lower overall. Again, there was no formal statistical test of these differences between these two groups of black students. There may have been a protective mental health effect for black students who had attended a largely non-white school. The analyses also showed that black students who had attended a school with a high percentage of white students (> 57%) in the 1994/95 school year had a higher depressive symptom rate at age 34 compared with white students who had attended schools with a similar racial composition. Both white and black depressive symptom reporting decreased over time, suggesting some beneficial health effect of attending a highly segregated school with a large percentage of white students. Importantly, part of this difference for black students was explained by earlier reports of perceived discrimination and lower school attachment. However, the gap in symptom reporting between white and black students became non-significant once respondent SES was controlled for at age 34. It is difficult, however, to assess what this meant for blacks of different SES backgrounds, as there was no statistical test for a potential interaction effect. These findings could suggest that the mental health effect related to school composition faded over time for high-SES blacks but not for low-SES blacks.

A more detailed examination of the results also suggested that for Hispanic and Asian students who had attended a predominantly white school, there was higher depressive symptom reporting at age 34 relative to the white students. However, this was a result of a more steeply declining reporting rate for white students over time. In fact, the reporting rates for Asian and Hispanic students who had attended a predominantly white school declined marginally, which suggested some minimal benefit to health.

In contrast, Hispanic and Asian students who attended a school in 1994/95 with a very low percentage of white students showed higher levels of depressive symptoms relative to white students who had attended a school with a similar composition. However, there did not appear to be any long-term negative effect. At age 34 symptom reporting for Asian students nearly converged with that of black and white students, although no specific significance test of the difference was reported. Symptom reporting for Hispanic students was lower relative to white and black students at age 34. Again, there was no reported significance test of this difference. Interestingly, there was a higher rate of symptom reporting among Hispanics and Asian students attending predominantly minority schools during the first wave of the survey that was not explained by feelings of discrimination or lack of school connectedness. The authors suggested in one of their earlier cross-sectional studies that this measure of discrimination might not have been sensitive enough to detect an effect related to race/ethnicity. It is possible, at least for most Hispanic students, that they simply attended schools that were predominantly Hispanic in composition.
The Walsemann et al. (2011) study included students enrolled in grades 9–12 at the time of the baseline survey. Depressive symptoms were measured via a nine-item scale that was summed at each of the four time points. School racial composition at the baseline was calculated as the percentage of non-Hispanic, white students. Self-identified race/ethnicity was aggregated to the school level and weighted so that the measure would be representative.

Students needed to have at least information on depressive symptoms at the first and last survey points. The majority (60%) had information at all four time points. The final sample included 10,350 respondents enrolled in 80 secondary schools. With these data a three-level linear growth modelling strategy was used, including random intercepts and slopes for individual students and schools. Models controlled for baseline student and school-level covariates. These included age, race/ethnicity, immigrant status, gender, family structure and SES, perceived discrimination, the degree of school attachment, school SES, urbanicity of the school, the school’s regional location and school type. A measure of adult SES from wave 4 was also included in the full-form models. Sampling variables instead of the weights were included in the analyses to produce unbiased estimates. Hence, all analyses were unweighted.

A fourth longitudinal study that examined the relationship between secondary school racial composition and later self-reported health in early adulthood produced mixed results in those 24–34 years old (Goosby & Walsemann, 2012). Findings suggested that black students who had attended schools with a large proportion of white students had lower self-rated health in early adulthood relative to whites. In contrast, Asian students seemed to benefit healthwise after attending schools with a similar very high white racial composition, one in excess of 70%. Asian students in schools with more ethnically diverse settings had lower self-rated health. However, the findings did not suggest that the later health status of Hispanic students or those with some other racial/ethnic identification were dependent on the white racial composition of the school.

Additional findings suggested an independent but negative effect of being treated unfairly by teachers on later health. However, this effect was completely mediated by the presence of strong parental support. Thus, the measures of discrimination did not appear to explain the negative health effects of black and Asian students. The authors again suggested that the measure might not have been sensitive enough, because it did not explicitly assess whether students felt that any perceived discrimination was racially motivated.

The Goosby and Walsemann (2012) study used data from the AddHealth waves 1 (1994/95) and 4 (2007/08). Thirteen thousand, five hundred and thirty-eight students from 132 schools were included in this analysis. School racial composition was measured as the percentage of white students at each school in 1994/95. During the first survey wave students were asked if they were treated fairly by teachers and whether other students at the school were prejudiced. Students were also asked to indicate how socially connected they felt in their respective schools and whether they received support from both parents. Statistical analyses controlled for age, gender, race/ethnicity, immigrant status, family structure, family SES, an assessment of self-rated health status, BMI, smoking behaviour, binge
drinking and drug use based on information collected at the first survey point. Composite measures of school SES were created from the student data and measures of school size, region location, type (public/private) and urban/rural location were included. The analyses did not use the provided sample weights but instead included control variables related to the sample respondents. The main analysis used two-level linear models. A sensitivity analyses was also run with a two-level logit modelling strategy.

Higher education: college selectivity and health

At least three studies have used American data to examine the relationship between college selectivity and various measures of health. One of the first studies to consider health effects also simultaneously tested the education quantity and education as a credential hypotheses (Ross & Mirowsky, 1999). Main results from reduced form models suggested that of the three indicators, the quantity of education had the strongest effect on self-assessed health (standardised $\beta = 0.15$) and physical functioning ($\beta = 0.121$). Surprisingly, the completion of a college degree had no significant association with either indicator of health status. Importantly, there was a smaller net association between college quality and the measures of health (self-assessed $\beta = 0.037$; physical functioning $\beta = 0.042$) as well as an interaction between years of education and greater college selectivity. Oddly, the interaction results were not presented. It was suggested that the health benefit of more education was modified by higher quality schooling.

Other important findings from the full-form models suggested that both the effects of years of schooling and college selectivity on physical functioning were mediated to some extent by psychological resources, but mostly by lifestyle factors. Both coefficients became statistically non-significant after lifestyle factors were included. In comparison, results from the full-form, self-assessed health models suggested that the effect of college selectivity only became non-significant once the controls for lifestyle factors were included. Prior controls for social psychological resources only marginally reduced the size of the estimate. However, the effect of years of schooling remained statistically significant after controls for both social psychological resources and lifestyle factors were added. Both coefficients were, however, nearly indistinguishable in magnitude. Regardless, the results were interpreted as supportive of the view that the quality of the school was less important than the quantity of schooling and that the effect of college selectivity was mostly mediated by lifestyle rather than through employment and economic conditions. In addition, findings related to college selectivity suggested that those who attended more selective schools were more likely to walk and exercise, and were less likely to be overweight or abstain from drinking. Attendance at a more prestigious college did not, however, translate into a lower probability of smoking, but that relationship is likely to have changed in more recent cohorts. This would be true of the general decline in smoking among the wealthier and the better educated.

The Ross and Mirowsky (1999) study used data from the 1995 Aging, Status, and the Sense of Control (ASOC) survey. It is a national telephone
probability sample of English-speaking adults in the USA. Individuals aged 60 and older were over-sampled. The sample was weighted so that it would be roughly representative of the US population. Full-form statistical models included total years of formal education, a dichotomous indicator for completion of a bachelor’s degree or higher, and an eight-category indicator for college selectivity (based on grade point average, the SAT/ACT scores of incoming freshman and the percentage of accepted applicants). In addition, there were controls for age, gender, marital status, parental education, race, work and economic conditions, social psychological resources and lifestyle indicators including BMI, drinking quantity and frequency, and smoking status.

Fletcher and Frisvold (2011) examined the relationship between the quality of higher education measured as college selectivity and a wider set of measures of health behaviour. These included the use of tobacco and marijuana, binge drinking, exercise, fast food consumption, and several different measures of body weight. Findings from the main analyses suggested that during college, students who attended higher quality (more selective) schools had a lower probability of tobacco (−6.5%) and marijuana (−3.8%) use, as well as a lower probability of obesity (−2.9%) and fast food consumption (−4.9%). After college attendance there was evidence that the benefits of higher quality schooling had a carry-over effect on health behaviours including a lower probability of smoking (−5.3%) and marijuana use (−3.3%), a lower probability of being overweight (−12.6%) or obese (−4.9%), lower BMI as well as a slightly increased probability of exercising (3.3%). The evidence on binge drinking, however, suggested that selectivity increased the probability of this behaviour (6.3%).

The study used nationally representative survey data obtained from the AddHealth school-based surveys (four waves). Nearly 15,000 students from wave 1 (1994/95) were followed over time until the final interview in 2008. Additional information on college selectivity was measured by a school’s national ranking using the median SAT score of the entrants. This was available for 4200 students who had enrolled in college at wave 3. Selective colleges were defined as those with median SAT scores in the top 25%. All other colleges were classified as non-selective. Controls for family background, student characteristics including high school grade point average, an ability measure, college expectations and health behaviours (binge drinking, smoking, marijuana use, and five weight-related measures) were available prior to the start of college.

A variety of modelling strategies were used because of substantial differences between selective and non-selective college attendees on many of the pre-college control variables. These included classmate comparisons and matching estimators in an attempt to address concerns about endogeneity, particularly unobservable variable bias. Models were adjusted for the school clustering and school fixed effects. Analyses were also run on the post college findings in an attempt to determine to what extent other factors like ability, the quantity of education, earnings and marital status had on the main results.

Findings from these other analyses suggested that the effects of college quality on smoking, binge drinking, marijuana use and weight were reduced to a small extent (one-fifth to one-third) by controls for education quantity (smoking/marijuana use),
income and marital status (binge drinking). Between 20% and 30% of the college quality obesity effect was explained by the quantity of schooling. A larger reduction in the quality effect of schooling (40–50%) was shown for fast food consumption and exercise differences at wave 4. Overall, this led the researchers to conclude that even after holding educational quantity, marital status and income constant, educational quality appeared to have the potential to improve population health via health behaviour nearly a decade after attendance. The authors speculated that direct health knowledge and college peer effects were important determinants of later health related behaviour.

Fletcher and Frisvold’s (2014) most recent longitudinal study focused on the long-term relationship between health behaviours and education quality. It was measured by selective college graduation in the USA during the 1960s. The main results from family fixed effects models suggested a substantial impact of education quality on the probability of being obese (15–18% lower) in older persons, with older being defined as individuals in their 60s. Similar findings were evident for older selective college graduates in terms of reduced BMI (4–6% lower). These results were robust to alternative modelling specifications and independent of marital status, income and occupational characteristics (e.g. insurance coverage and access to health care). These findings were viewed as evidence that health behaviours were influenced by higher education quality. The authors further speculated that the health advantage of higher quality education could be explained by increased health knowledge, increased productive efficiency, college peer effects on health behaviours and workplace policies available to those with higher quality education. However, similar to the findings of Ross and Mirowsky (1999) there were no consistent findings related to a reduction in the probability of smoking among those who graduated from more selective colleges.

The Fletcher and Frisvold (2014) study data came from the Wisconsin Longitudinal Study (WLS). It included a one-third random sample of the Wisconsin state high-school graduating class of 1957. The sample consisted of 10,317 students, followed until 2004. Data collection happened in 1957, 1964, 1975, 1992/93 and 2003/04. In 1977, 1993/94 and 2005–2007 additional data were also collected from randomly selected siblings. The sample was representative of white, non-Hispanic American men and women who completed high school. Prior to college attendance, information was collected on family economic conditions, socio-demographic characteristics including IQ, childhood health status, and health behaviours, in addition to information on the name of the degree granting college that was attended. Additional data on college selectivity (college ranks based on the incoming freshman class median SAT scores, grade point average and high school rank) were matched to the survey data. The final analytical sample included 3210 individuals who had complete information on BMI and current smoking status in 1993 or 2004, as well as information on the selectivity of their college. Family fixed effects models were run on a smaller set (n = 530) of discordant sibling pairs (where one graduated from a selective college and the other from a less selective school) in order to alleviate concerns about observable and unobservable background differences in the full sample.
Discussion

Taken together, evidence from the reviewed studies suggests that various attributes of the education system can influence health inequalities through the quality of the education available to children, adolescents and young adults. Findings suggested that several dimensions of school quality experienced throughout the life course directly affected health, or modified the effect of years of completed education on health. These included the pupil–teacher ratio, average teacher wages, and the length of the school year. There are both short and longer-term effects of school quality. The short-term effects are narrow and mixed while the later, long-term effects are more consistent, positive and apply to a broader set of health outcomes. However, not all measures of better school quality were related to better health. As noted earlier, smaller class size did not produce a beneficial health effect. It may have even produced marginally higher mortality in those exposed. Studies that attempted to assess why school quality produced beneficial health effects found that health behaviours and lifestyle factors were central. Occupational characteristics/choice and income also mediated the pathway between school quality and health. Net of these explanatory factors, residual effects were interpreted as the results of cognitive benefits and the uptake of better health information.

The two small, longitudinal studies using data from different RCTs assessed the health effects of exposure to high-quality pre-school. Both showed limited, long-term benefits. Given the advantages that treatment group subjects had in their educational attainment and income, and their relatively young age at the time their health outcomes were assessed, it is possible that additional beneficial effects of high-quality pre-school will emerge over time and will affect a wider set of health outcomes.

In contrast, the evidence from two longitudinal studies and a cross-sectional study on the effects of small class size on later adolescent/early adult health outcomes has suggested a null or negative effect. These findings are surprising given the disputed, but reasonably well-documented beneficial effects of small class size on academic achievement, college entrance exam completion, college attendance/completion, income and improvement in later non-cognitive outcomes, especially for lower-income children. Additionally, the cohort involved in the two longitudinal studies was young and had not yet developed major health problems. It is possible that later mortality and disability findings might be different in studies with a longer follow-up. Similarly, the cross-sectional study focused on the health effects of class size measured at a time when the students were older.

In these studies of the health effects of high-quality pre-school and small class size it would have been advantageous to have information on the quality of later schooling at different stages of the life course. Such data could have been used to assess whether exposure to higher quality education throughout the life course augmented the effect of earlier exposure to high-quality pre-school or smaller class size. Conversely, an assessment of whether later exposure to poor quality schooling muted or negated earlier potential beneficial health effects could have been done. A shortcoming of most of the reviewed studies was that health behaviours and outcomes were usually assessed at a relatively early age. Later life course studies would be more informative about longer-term effects and could offer evidence that contradicts or further clarifies null and counterintuitive findings. Databases that track subjects until mid and later life are clearly needed to do this.
The two longitudinal studies that included other measures of school quality, including lower pupil–teacher ratios, teacher wages and a longer school year, have uniformly suggested positive health effects much later in the life course for those exposed to higher quality schooling. Lower mortality was observed among synthetic birth cohorts and Black-American cohorts born when school quality improved dramatically. Better self-rated health and a lower probability of smoking and obesity were also observed among the Black-American cohorts much later in life. Importantly, one study also assessed how higher quality schooling produced better health. After controlling for the mediating effects of occupation, access to health insurance, income and marital status, it was concluded that the residual effect of school quality on lower mortality and obesity, and better self-rated health, was likely attributable to the cognitive benefits of better schooling. Higher earnings and occupational choice were also each partially responsible for these findings. There appeared to be no mediating social network effect.

It is notable that both studies found that mortality was lower in states with higher relative teacher pay. However, aggregate-level teacher pay is a crude measure of teacher quality. Including measures of accumulated teaching experience and pedagogical certification in addition to, or in place of indicators of teacher pay could improve understanding of the instructional quality and health relationship. In contrast to the study of the increased mortality effect of smaller class size, lower pupil–teacher ratios were consistently associated with lower individual and aggregate-level mortality. One study showed a direct effect while the other study’s findings suggested that a lower pupil–teacher ratio augmented the effect of additional schooling on lower individual-level mortality.

Definitive conclusions about the effects of school segregation on health are not warranted given the inconsistent findings in the four studies reviewed. The only consistent finding suggested that black students who had attended predominantly white schools had poorer health outcomes relative to white students. It appeared that at least one of these studies could have determined if this was likely to have been the case for both high and low-SES black students. This type of analysis was not done. However, for black female students in highly segregated English tracks, it appeared less likely that they adopted health damaging behaviours. It could be concluded tentatively that for black students in the USA, attending a predominantly non-white school was better for their health, but findings for the other racial/ethnic groups were too mixed to draw firm conclusions. The health effects of school segregation appeared to be highly dependent on the racial/ethnic group composition of the school in question, making it difficult to generalise about negative or positive health effects of high segregation in US schools.

Given these inconclusive findings in relation to health effects of the racial composition of schools, it is clear that such studies should be replicated with other data sets that include a larger number of students from other racial and ethnic groups. Any new study or replication should also include a classroom-level measure to assess whether a student is embedded within a majority/minority track. Longer-term follow-up is also needed to assess whether negative health consequences carry over into the later life course.
In contrast, the three studies that focused on the health effects of higher college selectivity had mostly consistent findings. They found support for positive health outcomes in both early adulthood and older age, after college attendance. These outcomes included better self-assessed health and physical functioning, a higher probability of exercising, a lower probability of being overweight or obese, lower BMI among study subjects and a lower probability of smoking marijuana. However, results related to smoking were inconsistent. Two of the studies found that higher college quality did not produce a lower probability of smoking later, while one study did find such a relationship. One study also found a higher probability of later binge drinking while another found a lower probability of abstaining from drinking. Binge drinking is likely to be damaging to health while abstaining from drinking has an ambiguous effect on health. One study also examined health behaviour during college attendance and found higher college quality was associated with a lower probability of tobacco and marijuana use, and a lower probability of obesity and fast food consumption.

Two of the college quality studies attempted to assess why beneficial health outcomes occurred among those exposed to higher quality schooling. One of the studies concluded that there was an interaction between college selectivity and years of schooling such that higher quality increased the benefit of the more selective education. The study further demonstrated that the direct effect of quality on health was mediated by lifestyle factors rather than through employment and economic conditions. The other study, after accounting for ability, the quantity of education, earnings and marital status in a stepwise fashion, concluded that college quality still had a direct effect on health. The third study, after employing similar controls, as well as occupational characteristics, came to the same conclusion about a direct, beneficial effect of college quality on health. All three studies suggested that these findings meant that direct health behaviour and lifestyle characteristics were at least partly responsible for the beneficial effect. The last two studies also speculated that college peer effects, increased health knowledge, and even workplace policies might have explained part of the health advantages of higher quality college attendance. This later group of studies is better situated to assess the longer-term health effects of school quality because they were based on longitudinal data and they employed a variety of different statistical approaches to strengthen causal inference.

The argument that higher levels of educational attainment produce an independent, positive effect on individual-level health is theoretically persuasive. Studies that have attempted to show a causal linkage have produced evidence that supports this interpretation (Lleras-Muney, 2005; Mazumder, 2007, 2008). Several of the studies reviewed in this paper suggested that the quality of education modified the effect of the quantity of education on health in ways that appeared to produce positive health outcomes over the life course. A number of these studies hypothesised that higher quality schooling enhanced cognitive ability. However, this has not been directly tested in any study that we are aware of. No school quality study to date has included a measure of a change in cognitive ability. To better assess this potential mechanism clearly requires such a measure. Unfortunately, most of the reviewed studies had no measure of baseline cognitive ability. The only exceptions were Fletcher and Frisvold (2011, 2014) studies and the Muennig et al. (2011a) study. Clearly, this is needed to reduce concern about potential selection effects. Evidence demonstrating that
cognitive ability can be modified by social structural influences in young children (Turkheimer et al., 2003; Tucker-Drob et al., 2011) further suggests that exposure to high-quality schooling itself is also likely to be a determinant of cognitive ability.

Considerable evidence supports the idea that a high-quality, compensatory educational approach, begun at a very early age and continued throughout the early life-course, is needed to help low-SES children from falling further behind their high-SES peers. Some have even concluded that without early, high-quality childhood education to increase the school readiness of low income children, education policies directed at improving the quality of primary and secondary education are unlikely to make much of a difference (Low et al., 2005). Moreover, better quality education in the early years is likely to lay the groundwork for greater educational attainment later in the life course (Frisvold & Golberstein, 2011). This could occur as a result of a more positive initial experience with learning, which could help to improve a child’s educational aspirations. Early high-quality schooling might also improve the returns to schooling as cognitive skills may be enhanced and or become more enduring as a function of such environments (Cutler & Lleras-Muney, 2010).

A final thought about policy is in order. If additional evidence about the positive relationship between characteristics of schooling and health is found, policy makers should think more about the costs of each. At present, the costs of a year of high-quality schooling, and greater quantities of schooling, appear to be high. When evaluated against the cost of health care—physicians, hospitals, pharmaceuticals—expenditures to increase the quality of schooling are likely to be more modest. Twelve years of high-quality schooling compared with 12 years of diabetes, or 5 years of dialysis, or one heart valve replacement, or one lung transplant make high-quality schooling look like a bargain.

The weight of the evidence we have presented suggests that more schooling modified by higher quality schooling can lead to improved health in individuals and in some groups of people. This could translate into reduced health-care system expenditures in addition to a diminution of emotional stress that illness brings to each of our families. The returns to society for better and more education, even without healthier outcomes for our citizens, appear to be substantial. Economists almost all agree that better and more education will have positive economic benefits. If the education–health link we reviewed is found to be consistent and significant with additional research, as we think it will prove to be, nations that improve the quality of their education systems will likely improve their economies, improve their citizen’s health and reduce their health-care costs. We urge researchers and policy makers to build databases that allow for further detailed investigation of these issues.

Acknowledgements

This research was partially funded by a grant from the European Commission Seventh Framework Program (FP7/2007-2013) No. 278511—DEMETRIQ (J. Mackenbach/M. Whitehead/O. Lundberg), and a Vetenskapsrådet grant: Swedish Research Council No. 2013-2139—Educational quality, birth characteristics, scholastic and employment outcomes for two generations of Swedish children (A.M. Garcy).
References


© 2018 British Educational Research Association


