

# Pathways between education and health: a causal modelling approach

Tarani Chandola<sup>1</sup>

*International Institute for Health, Well Being & Society, UCL, U.K.*

Paul Clarke and David Blane

*Imperial College London, U.K.*

and J.N. Morris

*London School of Hygiene and Tropical Medicine, U.K.*

**Summary.** The association of poor education and poor health has been consistently observed in a number of studies and in different countries. Thus far, studies examining the mechanisms underlying this association have only looked at a limited set of potential pathways. This study simultaneously examines six distinctive pathways, which have been hypothesized to link education and health and found support from previous studies. A causal analysis of education and health was performed using structural equation models. Data were used from six phases of the National Child Development Study, which is based on following up an initial sample of 17,416 children born in 1958. The association between education and health appears to be explained by a combination of mechanisms: adolescent health and adult health behaviours for men and women; adult social class among men; and parental social class among women. We conclude that improvements in population educational attainment may not automatically lead to improvements in population health, and that health policies for improving health and reducing health inequalities need to target specific causal pathways.

*Keywords:* Education, Health, Pathways, Structural Equation Modelling

## 1. Introduction

In 1992, the U.K. Department of Education first published league tables of GCSE pass rates in the different Local Education Authority (LEA) areas of England. The response of

---

<sup>1</sup>*Address for correspondence:* Tarani Chandola, Institute for Health and Well Being, Department of Epidemiology and Public Health, UCL, 1-19 Torrington Place, London WC1E 6BT, U.K.  
Email: [t.chandola@ucl.ac.uk](mailto:t.chandola@ucl.ac.uk)

one experienced epidemiologist (JNM) on viewing the league tables was “this list looks familiar”, because it closely resembled a list of the same areas ranked by mortality. An analysis of these results by Morris *et al.* (1996) confirmed this impression: at the LEA level, the correlation between the all-cause mortality rate and the proportion of children aged 15-16 years passing five or more GCSEs at grades A to C was  $-0.77$  for males, and  $-0.64$  for females. In addition, the correlation at the LEA level between the proportion of children passing five or more GCSEs and the Townsend index of deprivation was  $-0.89$ .

Blane *et al.* (1996) looked further at the relationship between GCSE pass rate, mortality and deprivation in the LEA areas. They found that deprivation explained much of the relationship between education and mortality, whereas little of the relationship between mortality and deprivation was explained by education. A naive interpretation of this finding would be that deprivation affects mortality, and that deprivation also plays some mediating role for the relationship between education and mortality. In reality, as the authors acknowledged, the conclusions that can be drawn from this analysis are extremely limited. Not only does the ecological fallacy preclude drawing reliable inferences about individuals from aggregated data, but any tentative causal interpretation of the results is prevented by the fact that pass rates measure educational attainment of adolescents, whereas mortality rates quantify deaths among (mainly) older individuals. However, the above illustration demonstrates the need to understand and quantify the effect of the pathways linking education and health and how these could contribute to the debate on the role of education in reducing health inequalities.

Studies of education and health have already been undertaken using disaggregated, individual-level data, particularly in economics (e.g. Berger and Leigh 1989; Feinstein 2002; Murasko 2003), epidemiology (e.g. Smith *et al.* 1998; White *et al.* 1999), and sociology (e.g. Ross and Wu 1995). More comprehensive reviews of the area can be found elsewhere (Blane 2003; Mirowsky and Ross 2003). In each of these studies, statistical techniques were used to estimate the direct effect of education on health, with proper adjustment for factors that could induce a spurious association between the two. From these studies, it was concluded that any explanation of the relationship in terms of a direct effect between education and health would likely be inadequate, and that the true effect is more likely the result of multiple mechanisms by which education influences intermediate factors that in turn influence adult health. However, only one study to date has explicitly considered multiple pathways in its analysis (Mirowsky and Ross 1998).

In his review of previous studies on education and health, Blane (2003) identified five mechanisms, or pathways, which found consistent theoretical and empirical support. Feinstein (2002) noted another important pathway which has found support in economics. In short, these six pathways are: 1) parents' socio-economic circumstances directly affect health and educational achievement; 2) educational achievement affects socio-economic status in later life which in turn affects health; 3) education affects receptivity to health messages which affects health; 4) education affects sense of control which affects health; 5) childhood cognitive ability mediates the education-health relationship; and 6) childhood health directly affects both health in later life and education.

In this paper, we aim to add to the debate on education and health in Britain by investigating all six of the hypothesised pathways using longitudinal data recorded on a single birth cohort. By considering all six pathways simultaneously, we can look at the wider effect of education on health through ‘indirect’ as well as direct effects. For an example of an indirect effect, take pathway 4 above; the indirect effect of education on health operates through sense of control, because education affects an individual’s sense of control in later life, which determines (in part) their adult health. In another context, Singh-Manoux *et al.* (2002) demonstrated how temporally distant effects such as education may be understated if indirect effects are ignored. Pearl (2001) stated that effective policy interventions typically require evidence for total and indirect effects, not just direct effects. Considering all six pathways together further allows us to consider the relative impact of each pathway, and to determine whether any of the pathways are spurious, that is, solely due to chance associations with the factors in other pathways.

We cannot use standard regression modelling methods because the causal sequence implied by the six pathways is too complicated, and estimation of indirect effects is not straightforward. As such, we follow Mirowsky and Ross (1998) and use causal modelling techniques to facilitate this analysis (e.g. Pearl 1995; Greenland 2000; Sobel 2000; Greenland and Brumback 2002). Causal models are a family of statistical techniques through which pathways can be explicitly modelled and tested. More specifically, we use structural equation models as implemented in the software package Mplus (Muthén and Muthén 2003).

The data for this study come from the National Child Development Study, which is described in more detail in Section 3 (Ferri 1993). It is a multi-disciplinary, on-going longitudinal study based on 17,416 individuals born in Great Britain in the week 3-9 March 1958. The study is multi-disciplinary and contains a rich variety of variables with which the key factors in the pathway model can be quantified.

The structure of this paper is as follows. In Section 2, the six pathways between education and health are described in detail. In Section 3, the National Child Development Study data and the variables from it used in the analysis are described. In Section 4, details about structural equation models as implemented in Mplus are given, and the analytical approach used in the paper described. The results are presented in Section 5, followed by discussion and some conclusions in Section 6.

## **2. Pathways Between Education and Health**

Blane (2003) reviewed the literature in which the links between education and health were investigated, and highlighted five plausible pathways for which evidence has been consistently found. In addition, we have added cognitive ability to this list (Feinstein 2002).

1. The socio-economic circumstances of the parents may independently influence their children's educational attainment and health. Higher parental social class is strongly associated with greater parental interest in the child's education and consequently, better educational attainment (Feinstein and Symons 1999). Lower parental social class (in childhood) has also been strongly associated with greater

morbidity in adulthood (Power and Hertzman 1997). The association between education and health could therefore arise from the effect of parental social class on educational attainment and adult health.

2. A person's educational attainment may influence their socio-economic circumstances later on in life which may, in turn, influence their adult health. Some studies have found the effect of education on adult health becomes non-significant after adjusting for occupational class and income (Dahl 1994; Smith *et al.* 1998). This suggests that the effect of education on health operates through a pathway where higher education leads to higher occupational attainment and income, which, in turn, increases the chances of better health. In contrast, other studies have found the effect of education on health remains strong even after controlling for occupational class and income (Winkleby *et al.* 1992). It thus remains debatable whether the effect of education on health can be entirely explained by its effect on income and occupation.
3. Education might affect a person's receptivity to health education messages (such as the dangers of smoking or the importance of health checks) which could have a beneficial influence on their health through health promoting behaviours and lifestyles (Fuchs 1979; Kenekel 1991). There is some evidence that greater education is associated with lower rates of smoking (Winkleby *et al.* 1992) and regular health checks (Coburn and Clyde 2002). However, nearly all studies report that even after adjusting for healthy behaviours and lifestyles, educational attainment remains significantly associated with health (Ross and Wu 1995).

4. Education may affect a person's sense of control over their life and this sense of control may affect their future health status. Education increases the sense of personal control by developing analytic and communication skills (Mirowsky and Ross 1998). The sense of personal control may also improve health through enhancing healthy behaviours, by controlling one's immediate addiction for future long term health benefits (Folkman 1984), while the lack of personal control may be a stressor with consequent adverse physiological consequences (Ross and Wu 1995).
5. Childhood cognitive ability (or intelligence) may also affect both educational attainment and adult health outcomes (Gottfredson 2004; Gottfredson and Deary 2004, Batty and Deary 2004). Cognitive ability in learning, reasoning and solving problems affects academic performance positively as well as job performance and functional literacy (Gottfredson 2004). Furthermore, intelligence predicts lower mortality rates from all causes, cardiovascular disease and lung cancer (Hart *et al.* 2003). Intelligence could enhance a person's care of their own health through better receptivity to health education messages (mechanism 3) by representing learning, reasoning and problem solving skills useful in preventing chronic disease and accidental injury and in completing complex treatment regimes (Gottfredson and Deary 2004). Although, this mechanism has been hypothesized in the literature, there have been few, if any studies that have empirically analysed this pathway in relation to all the other mechanisms underlying the association between education and health.

6. A child's health may independently influence their educational attainment (Isohanni *et al.* 2001) and their later, adult health (Jefferis *et al.* 2002). Ill health or disease during childhood limits educational attainment and predisposes individuals towards adult morbidity (Hotopf *et al.* 1999). However, the effect of childhood health on educational attainment and adult health has seldom been compared with the other mechanisms linking education and health.

No study to date has considered all six pathways simultaneously. The six pathways are shown graphically in Figure 1a. Factors mediating the pathways are shown as boxes or circles (the distinction between the two will become apparent in Section 4); single-headed arrows denote the causal effects of one factor on another; and non-causal associations are denoted by lines with double-headed arrows. To aid clarity, the two childhood cognitive ability variables (reading and mathematics) are represented by a single box, Cognitive Ability. To further aid clarity, the pathway diagram with only the causal effects of primary interest is shown in Figure 1b. The pathways excluded from Figure 1a are: the causal pathways from childhood cognitive ability, childhood social class and adolescent health to each of adult social class, sense of control and healthy behaviours. The rationale behind the inclusion of each variable in these diagrams is discussed in the next section.

### **3. National Child Development Study**

The National Child Development Study (NCDS) is a continuing, multi-disciplinary longitudinal study which takes as its subjects all those living in Great Britain who were born in the week between 3-9 March 1958 (Ferri 1993). Following the initial 1958

survey, called the Perinatal Mortality Survey (PMS), there have been six attempts to trace all members of the birth cohort and monitor their physical, educational and social development. These were carried out in 1965 (at age 7, called NCDS1), 1969 (age 11, NCDS2), 1974 (age 16, NCDS3), 1981 (age 23, NCDS4), 1991 (age 33, NCDS5) and 1999 (age 42, NCDS6). The NCDS is a unique, publicly available, data source containing measures on individual characteristics and background information from different stages of the life course, which is essential if we are to model the hypothesised mechanisms. More precisely, we need measures of the following groups: adult and adolescent health status, sense of control, healthy behaviour, education, childhood and adult social class, and childhood cognitive ability.

### *3.1. Adult health*

Current adult health is taken to be that at age 42. The NCDS has four variables related to adult health, self-rated health in general, self-rated health over the last 12 months, limiting (in terms of paid work) long-standing illness, and hospital admission (excluding accidents).

Rather than consider each of the four adult health variables separately, we chose instead to follow Mirowsky and Ross (1998) and Sadana *et al.* (2002) and analyse a single latent variable Adult Health, summarising the four variables. The latent variable approach is convenient because it allows multiple measures of the same characteristic to be included in the model. This approach may reduce potential bias from measurement error in the observed variables. For example, suppose there is a social comparison bias inherent in

self-rated health measurement (i.e. people comparing their health to others' within their own social group rather than in the general population), and memory bias in questions on hospital admissions; then the latent variable method will reduce measurement error bias if the social comparison and memory biases are in opposite directions. In Section 4, we describe how to specify the Adult Health latent variable using the four health variables in the measurement model component of a structural equation model.

The four adult health variables were recoded into binary variables with '0' denoting poor health and '1' denoting good health. If the person rated their health in general and over the last 12 months as fair or poor, had a limiting long-standing illness or hospital admission, they were said to have poor health.

### 3.2. *Adolescent health*

As with adult health, we chose to use a latent variable, Adolescent Health, to represent adolescent health status at age 16 because single item measures may be subject to bias (Power *et al.* 1990). The four NCDS variables are: General Practitioner (GP)/health centre attendance; hospital admissions in the last year for at least one night; frequency of school absence (due to ill health) in the last year; and disability. The four Adolescent Health variables were then recoded into binary variables with '0' denoting poor health and '1' denoting good health. If the child attended a GP/health centre four or more times in the previous year, had been admitted to hospital for at least one night in the previous year, had a disability or handicap, or were absent from their school for more than a month in the previous year due to ill health, the adolescent was said to have poor health. A

measure of adolescent health was chosen over adolescent health because of its stronger association with educational attainment.

### *3.3. Sense of control*

The NCDS has three variables relating to Sense of Control, at age 33, which are combined into a single latent variable for modelling purposes (see Section 4).

Respondents were asked whether they had free choice and control over their lives, whether they got what they wanted out of life, and if they could run their lives as they wanted to (Ford *et al.* 2004). The (binary) questions were coded such that ‘0’ denoted control and ‘1’ denoted lack of control.

### *3.4. Healthy Behaviours*

There are three measurements relating to Healthy Behaviours at age 42 in the NCDS: regular exercise, fruit consumption, and smoking. These variables were recoded into binary variables with ‘0’ denoting unhealthy behaviours and ‘1’ denoting healthy behaviours. As with the preceding measures, Healthy Behaviours is included in the model as a latent variable (see Section 4).

### *3.5. Education, social class, and childhood cognitive ability*

Education at age 23 was measured by a six category ordinal measure of increasing academic or vocational qualifications from no qualifications to degree level or higher. Childhood Social Class (father’s social class at age 7, supplemented by father’s social class at birth if missing) and Adult Social Class (at age 33) were measured by the

Registrar General's social classification, an ordinal measure ranging from 1 (professional occupations) to 6 (manual unskilled occupations). Childhood cognitive ability (at age 7) was measured by maths and reading test scores (a higher score indicates greater ability). Ability tests at age 7 were chosen over intelligence tests (measured at age 11 in the NCDS) to reduce the potential for the confounding of family and school effects on intelligence.

#### **4. Data Analysis Using Structural Equation Models**

##### *4.1. The motivation for using structural equation models*

Most previous studies of education and health were based on regression models. For example, Smith *et al.* (1998) used proportional hazards regression to compare disease risk predictions using education groups with those using occupational status, adjusted for confounding bias; Berger and Leigh (1989) used instrumental variable regression models to estimate the direct effect of education on health while adjusting for a) endogeneity due to unobserved/unobservable factors influencing both education and health, and b) reverse causality whereby childhood health determines educational attainment and induces a spurious association between education and current health.

The focus of the studies above was on estimating the direct effect of education on health. The study by Ross and Wu (1995) is more similar to ours, in that it looked to establish the hypothesis that education determines health through work and economic conditions in later life, which in turn give rise to better social-psychological resources, which in turn results in better health-related lifestyle behaviours and so better health. Their analysis

consisted of fitting four regression models of increasing complexity, with each model including a set of socio-demographic variables to adjust for confounding bias.

Standard regression model approaches can be used to consider direct effects mediated by a single pathway, such as that considered by Ross and Wu (1995). However, we wish to consider six pathways simultaneously, and to estimate indirect and direct effects of education on health. To do this, we consider a more advanced technique, namely, structural equation modelling. Mirowsky and Ross (1998) have also used structural equation models to test three pathways specific to the human capital hypothesis on data from a cross-sectional sample of the US population.

#### *4.2. Structural equation models*

Structural equation models (SEMs) are an extension of standard regression models through which multivariate outcomes and latent, or unobserved, variables can be modelled. SEMs are more appropriate for this application than alternative causal modelling techniques such as graphical modelling because they permit specification of ‘measurement models’. Measurement models are suitable for observed outcomes which are believed to be imperfect measures of an unobserved (or unobservable) underlying construct (e.g., the constructs Sense of Control and Healthy Behaviours). We give a brief overview of structural equation models here, but leave the reader to look elsewhere for further details (e.g. Muthén 2003).

The general SEM has two components, a ‘measurement’ model and a ‘structural’ model, the standard notation for which is

$$z_i = \nu + \Lambda \eta_i + Kx_i + \varepsilon_i \quad (1)$$

$$\eta_i = \alpha + B\eta_i + \Gamma x_i + \zeta_i \quad (2)$$

where  $i$  indexes sample individuals,  $z_i$  is a vector of latent outcome variables,  $\eta_i$  is a vector of latent variables,  $x_i$  is a vector of background exogenous variables,  $\varepsilon_i$  and  $\zeta_i$  are normally distributed error terms,  $\nu$ ,  $\alpha$ ,  $\Lambda$ ,  $B$ ,  $K$  and  $\Gamma$  are the model parameters; model (1) is the measurement model, and model (2) is the structural model. We give details below of how the SEM for the pathway diagrams in Figure 1 is specified.

#### *Latent outcome variables and the measurement model*

The variables denoted by circles in Figure 1 are latent variables  $\eta_i$  (Adolescent Health, Sense of Control, Healthy Behaviours, Adult Health), with squares denoting either exogenous background variables  $x_i$  (Cognitive Reading Score, Cognitive Mathematics Score, Childhood Social Class) or endogenous outcome variables  $z_i$  (Education, Adult Social Class). Each circle represents a latent variable and its measurement model. For example, Healthy Behaviours is a latent variable with a measurement model based on three ‘items’, the observed variables measuring exercise, fruit consumption, and smoking. The observed items are dichotomous and were excluded from Figure 1 to aid clarity, but are included in the model itself as latent outcome variables  $z_i$ . A latent outcome variable is equal to an observed outcome variable if the observed variable can be treated as continuous (i.e., it has a ratio or interval scale). If the observed outcome variable  $y_i$  is dichotomous or ordinal, it is represented in the model by a continuous latent outcome

variable. In the case of dichotomous outcome variables, the relationship between the observed and latent outcome variables is given by

$$y_i = \begin{cases} 1 & \text{if } z_i > 0 \\ 0 & \text{if } z_i \leq 0 \end{cases}$$

where  $y_i$  is one of the dichotomised variables in the model such as exercise.

The model for Figure 1 consists of measurement models for each of the four latent variables, Adult Health, Adolescent Health, Sense of Control, and Healthy Behaviours.

For example, the measurement model for Adult Health is given by four equations

$$\begin{aligned} z_i^{GSRH} &= \nu_1 + \lambda_1 \eta_i^{AH} + \varepsilon_{1i} \\ z_i^{SRH12} &= \nu_2 + \lambda_2 \eta_i^{AH} + \varepsilon_{2i} \\ z_i^{LLSI} &= \nu_3 + \lambda_3 \eta_i^{AH} + \varepsilon_{3i} \\ z_i^{AHA} &= \nu_4 + \lambda_4 \eta_i^{AH} + \varepsilon_{4i} \end{aligned}$$

where the latent variables is  $\eta_i^{AH}$  = Adult Health, and the latent outcome variables are:

$z_i^{GSRH}$  = Self-rated Health (General),  $z_i^{SRH12}$  = Self-rated Health (last 12 months),  $z_i^{LLSI}$  = Limiting Long-standing Illness, and  $z_i^{AHA}$  = Adult Hospital Admissions. The  $\varepsilon$ -terms are independent normal errors, the coefficients enter the appropriate matrices in model (1), and the  $\lambda$ -parameters are referred to as ‘factor loadings’. A similar set of equations determines the measurement model for the other three variables.

### *Structural model*

It is the parameters of model (2), the structural model, which are of primary interest in the analysis. More specifically, the  $B$  matrix contains the regression coefficients  $\beta$  of each variable on the others, with  $\beta = 0$  if the two are independent. The structural model

involves a predictor equation for each variable in Figure 1 with at least one single-headed arrow entering it. The system of equations corresponding to Figure 1b is:

$$\begin{aligned}
z_i^{EDU} &= \alpha_1 + \beta_{11}\eta_i^{TH} + \gamma_{11}x_i^{CSC} + \gamma_{12}x_i^{COGR} + \gamma_{13}x_i^{COGM} + \zeta_{1i} \\
z_i^{ASC} &= \alpha_2 + \beta_{21}\eta_i^{TH} + \beta_{22}\eta_i^{SOC} + \beta_{23}z_i^{EDU} + \gamma_{21}x_i^{CSC} + \gamma_{22}x_i^{COGR} + \gamma_{23}x_i^{COGM} + \zeta_{2i} \\
\eta_i^{SOC} &= \alpha_3 + \beta_{31}\eta_i^{TH} + \beta_{33}z_i^{EDU} + \gamma_{31}x_i^{CSC} + \gamma_{32}x_i^{COGR} + \gamma_{33}x_i^{COGM} + \zeta_{3i} \\
\eta_i^{HB} &= \alpha_4 + \beta_{41}\eta_i^{TH} + \beta_{42}\eta_i^{SOC} + \beta_{43}z_i^{EDU} + \beta_{44}z_i^{ASC} + \gamma_{41}x_i^{CSC} + \gamma_{42}x_i^{COGR} + \gamma_{43}x_i^{COGM} + \zeta_{4i} \\
\eta_i^{AH} &= \alpha_5 + \beta_{51}\eta_i^{TH} + \beta_{52}\eta_i^{SOC} + \beta_{53}z_i^{EDU} + \beta_{54}z_i^{ASC} + \beta_{55}\eta_i^{HB} \\
&\quad + \gamma_{51}x_i^{CSC} + \gamma_{52}x_i^{COGR} + \gamma_{53}x_i^{COGM} + \zeta_{5i}
\end{aligned}$$

where  $\eta_i^{TH}$  = Adolescent Health,  $\eta_i^{SOC}$  = Sense of Control, and  $\eta_i^{HB}$  = Health Behaviours,

$\eta_i^{AH}$  = Adult Health; and the exogenous background variables are  $x_i^{COGR}$  = Cognitive

Reading Score and  $x_i^{COGM}$  = Cognitive Mathematics Score, and  $x_i^{CSC}$  = Childhood Social

Class. The other variables are, in the parlance of SEMs, technically latent outcome

variables, and as such should be included in the measurement model. However, their role

in this analysis is as endogenous causal factors, and so for presentation purposes we

include them in the structural model. These are the ordinal variables  $z_i^{EDU}$  = Education,

and  $z_i^{ASC}$  = Adult Social Class.

#### 4.3. Parameter estimation and missing data

Parameter estimation is performed by maximum likelihood (ML) estimation in Mplus. If

data on the outcomes are missing, Mplus can perform ‘full information’ ML estimation

provided that a) the data can be assumed to be either missing completely at random

(MCAR) or missing at random (MAR) (Little and Rubin 2002); and b) the latent

outcomes and latent variables are normally distributed. Assumption b) corresponds to the standard SEM assumption and, as such, is not restrictive. However, in most analyses the MCAR assumption is unlikely to hold; the MAR assumption is more realistic if factors strongly associated with non-response or drop out are included in the model. In the NCDS data, out of the 17,416 babies in the initial birth cohort, 10,979 adults remained in the cohort at age 42, which corresponds to an attrition rate of around 37 percent. Hawkes and Plewis (2004) analysed patterns of attrition among the NCDS cohort. The results suggest that being born in Wales, having a mother with lower education and low birth weight significantly predict drop out. Including these variables in the model as covariates makes the MAR assumption more plausible. As such, they are included in each of the following structural equation models as background covariates (although they do not appear in the diagrams). If covariates are missing, similar arguments to above hold provided the additional assumption is made that c) the distribution of the missing covariate, given the other variables, is normal.

#### *4.4. Analytical strategy*

The analysis presented in Section 5.2 is broken into three parts:

- 1) Bivariate analysis. Estimate the bivariate associations between Adult Health and Education, and between Adult Health and each of the variables in Figure 1. This is performed using a series of SEMs, one for each pair of variables. The first SEM corresponds to a pathway diagram with a single-headed arrow drawn from Education to Adult Health; the second model corresponds to a pathway diagram with a single-headed arrow Cognitive Reading Score to Adult Health; and so on. This first part can

be thought of as a series of bivariate regression analyses and is used to establish the strength of the observed associations between the variables.

- 2) Adjusted Bivariate Analysis. Estimate these associations adjusting each for the other variables in Figure 1 in turn. This is performed using another series of SEMs. For example, to consider the effect of Education on Adult Health adjusted for Cognitive Reading Score, an SEM is used with double-headed arrows between Education and Cognitive Reading Score, and single-headed arrows from Education to Adult Health, and Cognitive Reading Score to Adult Health. This second part can be thought of as a series of adjusted bivariate regressions to estimate the adjusted effect of education.
- 3) Causal Analysis. The full set of direct and indirect effects of education on health were calculated using the full SEM specified in Section 4.3.

The analyses were performed separately for men and women using MPlus. Full information ML estimation was used to handle missing data.

## **5. Results**

### *5.1. Measurement model*

Four measurement models for the latent variables Adult Health, Adolescent Health, Sense of Control and Healthy Behaviours were specified in the analysis. The factor loadings (the unstandardised probit regression coefficients) are displayed in Table 1. The goodness of fit statistics for Adult Health and Adolescent Health indicate that the measurement models fit well. Hu and Bentler (1999) suggest the following fit index cut off value guide for good models with continuous and categorical outcomes: Tucker Lewis Index  $> 0.95$ , Comparative Fit Index  $> 0.95$ , and the Root Mean Squared Error of Approximation  $<$

0.06. Goodness of fit statistics could not be computed for the latent variable Sense of Control and Healthy Behaviours as there were only three indicators. However, each of factor loadings is significant, suggesting that there is common variance between all three indicators for both latent variables.

## 5.2. *Structural equation modelling*

### 5.2.1. *Men*

#### *Bivariate analyses*

The associations between Adult Health and Education, and between Adult Health and the other variables in Figure 1 are shown in the first column of Table 2. Bold numbers denote statistical significance at the 5% level. The first column displays the (unstandardised) coefficients of the regression of Education and Education-squared on Adult Health; the regression of Cognitive Reading Score on Adult Health; and so on. The Education-squared term was added because the effect of Education on Adult Health is non-linear. All of the variables were significantly associated with Adult Health, and all of these associations were in the expected direction; for example, men with higher educational qualifications at age 23 had better health at age 42. Those with higher scores on the arithmetic and reading tests (i.e. who had higher cognitive function) had better adult health. Those whose father's had worked in manual social class occupations had poorer adult health. Those with better adolescent health had better adult health. Men with a lower sense of control at age 33 had poorer health. Those who worked in manual social class occupations at age 33 had poorer adult health. Men with healthier behaviours at age 42 had better adult health.

### *Adjusted bivariate analyses*

The next six columns of Table 2 contain the adjusted associations between Education and Adult Health, adjusted for each of the other variables separately. Adjusting for Cognitive Ability, the association between Education (both the linear Education and squared Education-squared terms) and Adult Health is reduced slightly from 0.34 to 0.14 (linear) and  $-0.01$  to  $-0.00$  (squared). Similarly, the adjusting the Education-Adult Health association for Childhood Social Class reduces it from 0.34 to 0.17 (linear)  $-0.01$  to  $-0.00$  (squared); and ditto for Adult Social Class, Adolescent Health and Healthy Behaviours. The largest reduction in the linear effect of education on Adult Health occurs after adjustment for Adolescent Health and also adjustment for Healthy Behaviours.

### *Causal Analysis*

When fitting the full SEM for Figure 1 to the NCDS data, the direct linear effect of Education on Adult Health changes direction and is negative; that is, men with higher qualifications have poorer adult health (although this negative association is not statistically significant). However, the indirect linear effect of education on adult health (through its effect on adult social class, control and health behaviours) is positive and significant. The overall total linear effect of education on Adult Health (which combines the direct and indirect linear effects) is also positive and statistically significant.

Adolescent Health (at age 16), Adult Social Class (at age 33), Sense of Control (at age 33) and Healthy Behaviours (at age 42) have direct effects on Adult Health, with the first three also having significant indirect effects. Among the six mechanisms analysed, those

working through adolescent health, adult social class, sense of control and healthy behaviours had significant total effects on adult health. In contrast, cognition and father's social class did not have significant total effects on adult health.

### *5.2.2. Women*

A similar picture was obtained when analysing the pathways between education and health for women (Table 3). In the bivariate analyses, the association of Education with Adult Health was positive and non-linear. The associations with Adult Health of all the other variables were statistically significant and in the expected directions. In the adjusted bivariate analyses, the linear effect of Education on Adult Health was somewhat reduced when adjusted for any of the other variables.

In the causal analyses, the direct linear effect of Education on Adult Health changed direction and was negative. However, as with the men, the indirect linear effect and total linear effect of Education on Adult Health is positive. Healthy Behaviours had a significant direct effect on adult health. Adolescent Health, Sense of Control and Cognitive Reading Score also had significant direct effects on Adult Health, although only the former two affected Adult Health through indirect pathways as well. Childhood Social Class had indirect and total effects on adult health, although there was no direct effect of Childhood Social Class on Adult Health.

### *5.3. Summary of findings*

The main (standardised) pathway coefficients for the pathway models used to produce Tables 2 and 3 are displayed in Figures 2 and 3. Among both men and women, each of the childhood variables (Adolescent Health, Childhood Social Class, and Cognitive Ability) affect Education, which in turn affects Adult Social Class, Sense of Control, then Healthy Behaviours, and then Adult Health. The direct linear effect of Education on Adult Health is negative among both men and women, although it is statistically significant only among women. Adult Social Class and Sense of Control have direct effects on Adult Health among men but only the latter is significantly associated with Adult Health among women. Adolescent Health and Healthy Behaviours are the strongest influences on adult health among both men and women. (It is important to remember that not all the pathways and correlations in the models in Tables 2 and 3 are shown in Figures 2 and 3. However, the estimates of the hidden associations are displayed in the table below the figure.)

In summary, the association between Education and Adult Health appears to be explained by a combination of mechanisms. Adolescent Health is important for the health of men and women with both direct and indirect effects. Health Behaviours (measured by a latent variable consisting of smoking, exercise and diet factors) are also an important mechanism for men and women. Adult Social Class has direct and indirect effects on Adult Health among men, while Childhood social class has an indirect effect on women's Adult Health. Men and women with greater educational attainment have a greater Sense of Control, which in turn leads to better Adult Health. Although Cognitive Reading Score among girls was associated with better Adult Health, it did not have indirect effects on

Adult Health. Neither measures of Cognitive Ability explained the association of Education with Adult Health.

## **6. Discussion**

The results of this study are novel and represent a significant advance in the understanding of the association between education and health. In previous studies, it was found that education had a strong and positive direct effect on adult health, even after adjusting for explanatory variables. In our analysis, the ‘direct’ association between educational qualifications and adult health did not behave in this way. In men, the direct effect of education on adult health was not significant, while in women the direct effect was negative. In other words, if we compare women of the same childhood social class, cognitive ability, adolescent health, adult social class, sense of control and health behaviours, the health of women with higher qualifications is likely to be worse than for women with fewer qualifications. Other studies have found similar results (e.g. Singh-Manoux *et al.* 2002). As such, we found little evidence for the hypothesis that increases in educational qualifications ‘directly’ translates into better health.

It is important to emphasise that the behaviour of the direct effect of education does not imply education does not have a positive effect on health. A more realistic measure of the effect of education is its ‘total effect’, which can be estimated using our structural equation model and includes the effect of education on social class, the sense of control and health behaviours). The total effect of education was found to be positive for men and women, and it can be decomposed into the contribution from each of the underlying

pathways. We found that adolescent health, sense of control and adult healthy behaviours play an important role in explaining the education-health association, among both men and women. Furthermore, among men, adult social class also has some role in explaining the association, whereas among women, childhood social class is another mechanism. The elucidation of these pathways underlying the association between education and health have not been analysed and quantified in previous research using British data.

Another novel result is that it has been empirically demonstrated that childhood cognitive ability does not explain the association of education with adult health. Although reading ability among girls was associated with better adult health in the full model, it did not have indirect effects on adult health. The Scottish Mental Health Survey of 1932 found that adjusting for adult occupational social class attenuated the effects of intelligence on morbidity and mortality (Hart *et al.* 2003). In this study, the association between ability and adult health among boys was reduced considerably when childhood and adult social class, adolescent health, sense of control and health behaviours accounted for. Others have found that childhood ability, whether measured by reading/maths tests or intelligence tests, declines in predicting educational attainment relative to family background (Galindo-Rueda and Vignoles 2004). As neither the direct nor indirect effects of childhood cognitive ability on adult health was strong, the results suggest little evidence for the hypothesis that intelligence is the ‘fundamental cause’ of social inequalities in health (Gottfredson 2004).

A greater perception of control at age 33 was related to higher educational attainment, healthier behaviours and better adult health. This sense of personal control has been discussed in the literature in a number of related forms with different ways, such as internal locus of control (Rotter 1966), mastery (Pearlin *et al.* 1981), personal efficacy (Downey and Moen 1987), and personal autonomy (Seeman and Seeman 1983). The sense of control increases with socioeconomic status and appears stable until old age (Mirowsky 1995). Perceptions of control in adulthood appear to originate from higher parental education, higher cognitive ability and higher educational attainment (Lewis *et al.* 1999). These results are similar to the ones found in this study- a lower sense of control was associated with lower reading ability (at age 7), lower educational attainment (at age 23) and lower social class (at age 33). It is possible that children with a greater sense of control go on to do better at school and have better educational, occupational and health outcomes. The NCDS data did not have any measures of sense of control in childhood or adolescence and so we could not look at this hypothesis directly. However, other research has found that having a low sense of control in adolescence is not associated with lower educational attainment (Lewis *et al.* 1999), suggesting that modelling the effect education (at age 23) on adult health (at age 42) through sense of control (at age 33) is theoretically valid.

Good health in adolescence predicted higher educational attainment and better adult health. Although, in general, adolescents are biologically robust, with low morbidity and mortality, poor self-reported health at age 15 is strongly associated with medically unexplained physical symptoms at age 36 (Hotopf *et al.* 1999). Children with headaches

are at an increased risk of recurring headache, multiple physical symptoms and psychiatric morbidity in adulthood (Fearon and Hotopf 2001). The strong associations between adolescent health and adult health in this study are in line with these other findings.

Social class in adulthood had a direct effect on the health of adult men but not women. This gender difference is not unexpected- several other studies have found that when women are assigned to a class on the basis of their own occupation, the association between social class and health is not strong (Chandola 1998). On the other hand, classifying women by their head of household's class results in stronger associations between social class and health. The indirect effect of women's father's social class (at age 7) on their adult health in this study also suggests that household social class is important for the health of women.

Although the health behaviours examined in this study were limited (self-reported exercise, fruit consumption and smoking) and crudely measured, healthy behaviours were strongly associated with better adult health. Furthermore, these behaviours were influenced by education, social class and perceptions of control- those with higher educational qualifications, in non-manual occupations and with a higher sense of control had healthier behaviours. This suggests that in order to bring about improvements in health behaviours in the population, we need to understand how people's behaviours are influenced by social and psychosocial factors throughout their life.

The research literature on education and health reviewed in this paper has largely come from the disciplines of sociology and epidemiology. In economics, the human capital theory of health developed by Grossman (1972) also suggests a causal relationship from education to health, while conceding that the direction of causality may be reversed or that a ‘third variable’ that determines both education and health, drives their correlation. The ‘reverse causality’ hypothesis is incorporated into this study by including a latent variable measure of adolescent health in the analysis. The ‘third variable’ hypothesis was also examined by including measures of childhood and adult social class, ability, sense of control, and health behaviours. However, there may be other variables that were not included in this analysis that could drive the association between education and health. Such variables from economics literature on the association between education and health includes ‘future orientation’ (Fuchs 1982) (that is, people who have a high degree of time preference for the future attend school for longer periods and make greater investments in their health) and ‘psychological capital’ (Feinstein and Symons 1999; Murasko 2003) - including locus of control, self-esteem, hostility and anxiety. Some of this psychological capital may be captured in our analysis through the variable ‘sense of control’. It is possible that other variables, not included in our analysis, cause the relationship between education and health, although the set of (theoretically derived) variables examined in our analysis do a good job already in explaining the direct effects between education and health.

The statistical methods used in this paper were key to the extra insights we have gained from this analysis. The use of measurement (latent variable) models for adult and

childhood health and health behaviours also enable better measures of these concepts by (potentially) reducing biases inherent in single item measures. Mirowsky and Ross (1998) also used structural equation models to address the education-health question, using cross-sectional data from the U.S. As such, our use of prospective data from the NCDS offers an improvement on their study commensurate with the benefits of prospective over retrospective measures. Against this, there is considerable drop out from the original cohort. However, the size of the sample responding in adulthood is adequate for the study of adult health (Wadsworth *et al.* 2003), and we have used estimation techniques which protect against non-response bias under the ‘missing at random’ assumption (Little and Rubin 2002). Finally with regards to methodology, it is important to note that we do not claim to have established the fundamental true cause of how education affects adult health, despite the ‘causal analysis’ tag. Rather, we have taken the most widely believed theories on how education relates to health, and assessed the empirical evidence for each.

The policy implications of this study are clear. Although a greater proportion of the population today are acquiring degree level educational qualifications, this may not ‘directly’ or automatically lead to better adult health. As we discussed at the start of this section, when we compare women of the same childhood social class, cognitive ability, adolescent health, adult social class, sense of control and health behaviours, the ‘direct effect’ of education on adult health was found to be negative; however, the total effect (which includes the effect of education on health through social class, the sense of control and health behaviours) is positive. This distinction is extremely important with regards to policy interventions. First, policies designed to improve adult health by improving

educational qualifications must be carefully implemented. If a policy adversely affects any of the intermediate pathways while improving education, it will fail, or have far less of an impact than was expected. Second, policies can be aimed at specific pathways rather than directly at education. Third, it may be that intermediate pathways are not stable and constant over time; in other words, exogenous processes change over time and affect the distribution among key factors in the pathway. For example, as more women enter the labour force and are in full time employment, the effect of adult social class on women's health may increase from the estimated effect in this particular cohort (social class in this analysis was measured in 1981). As more people obtain degree level qualifications, the returns to higher education in terms of occupational or psychosocial benefits may change. An examination of these pathways in a younger cohort who have experienced such changes may throw some light on whether changes in the distribution of these 'risk factors' for adult health affect the association between education and health.

Apart from providing evidence on what types of policies may reduce the association between education and health, this study also suggests when policy interventions are most useful. There are high returns associated with interventions in childhood and adolescence, improvements in adolescent health result in better educational and adult health outcomes. Interventions on encouraging healthy behaviours need to focus on occupational and psychosocial factors in adult life. Encourage girls from manual working class backgrounds to stay on in higher education could also have beneficial health effects.

There is little doubt of a causal relationship between education and health. However, the pathways underlying this association have been imperfectly understood. The analysis of these causal pathways in this study has shown that adolescent health, healthy behaviours and the sense of control contribute substantially to explaining this association.

Furthermore, adult social class among men, and childhood social class among women also contribute to explaining the association between education and health. Policies aimed at improving adult population health and reducing social inequalities in health need to focus in on the causal pathways highlighted in this study.

**Acknowledgements:**

The data and tabulations used in this publication were made available through The ESRC Data Archive. The data were originally collected by the National Birthday Trust Fund, Butler, N. (Perinatal Mortality Study), the National Children's Bureau (sweeps 1-4), City University. Social Statistics Research Unit (sweep 5), and the Joint Centre for Longitudinal Research which comprises the Centre for Longitudinal Studies, Institute of Education, University of London; the International Centre for Health and Society (ICHS), University College Medical School, London; and the National Centre for Social Research (Natcen)- sweep 6. Natcen was responsible for: assisting in the development of instrumentation, development of CAPI programs, conduct of fieldwork, initial data coding and editing, and some documentation. The Office for National Statistics coordinated funding on behalf of various UK Government departments. Neither the original collectors of the data nor the Data Archive bear any responsibility for the analyses or interpretations presented here.

The authors would like to thank Mel Bartley, Fiona Steele and Archana Singh-Manoux for their helpful comments on earlier drafts. TC and PC were funded by an Economic and Social Research Council grant (RES-000-22-0290).

## References

- Batty, D. and Deary, I. (2004) Early life intelligence and adult health. *Brit. Med. J.*, **329**, 585-586.
- Beger, M.C. and Leigh, J.P. (1989) Schooling, self-selection, and health. *J. Human Resour.*, **24**, 433-455.
- Blane, D. (2003) Commentary: Explanations of the difference in mortality risk between different educational groups. *Int. J. Epidemiol.*, **32**, 355-356.
- Blane, D., White, I. and Morris, J. (1996) Education, social circumstances and mortality. In: Blane, D., Brunner, E., Wilkinson, R. (eds.), *Health and Social Organization*. London: Routledge.
- Chandola, T. (1998) Social inequality in coronary heart disease: a comparison of occupational classifications. *Soc. Sci. Med.*, **47**, 525-533.
- Coburn, D. and Clyde, R. (2002) Socioeconomic status and preventative health behavior. *J. Health Soc. Behavior*, **15**, 67-78.
- Dahl, E. (1994) Social inequalities in ill health: the significance of occupational status, education and income – results from a Norwegian study. *Sociol. Health Ill.*, **16**, 644-667.
- Downey, G. and Moen, P. (1987) Personal efficacy, income, and family transitions: a longitudinal study of women heading households. *J. Health. Soc. Behav.*, **28**, 320-333.
- Fearon, P. and Hotopf, M. (2001) Relation between headache in childhood and physical and psychiatric symptoms in adulthood: national birth cohort study. *Brit. Med. J.*, **322**, 1145.

- Feinstein, L. (2002) Quantitative estimates of the social benefits of learning, 2: Health (depression and obesity). Wider Benefits of Learning Research Report No. 6. London: Centre for Research on the Wider Benefits of Learning.
- Feinstein, L. and Symons, J. (1999) Attainment in secondary schools. *Oxford Econ. Pap.*, **51**, 300-321.
- Ferri, E. (1993) Life at 33: the fifth follow-up of the National Child Development Study. National Children's Bureau, City University, Economic and Social Research Council.
- Folkman, S. (1984) Personal control and stress and coping processes: a theoretical analysis. *J. Pers. Soc. Psychol.*, **86**, 839-852.
- Ford, R., Marsh, A. and Finlayson, L. (2004) What happens to lone parents. Department of Social Security Research Report No. 77. London: HMSO.
- Fuchs, V. (1979) Economics, health and post industrial society. *Milbank Q.*, **57**, 153-182.
- Fuchs, V.R. (1982) Time preference and health: an exploratory study. In: Fuchs, V.R. (ed.), *Economic Aspects of Health*. Chicago: University of Chicago Press.
- Galindo-Rueda, F. and Vignoles, A. (2004) The Declining Relative Importance of Ability in Predicting Educational Attainment. *J. Hum. Res.* (in press).
- Gottfredson, L.S. (2004) Intelligence: is it the epidemiologists' elusive "fundamental cause" of social class inequalities in health? *J. Pers. Soc. Psychol.*, **86**, 174-199.
- Gottfredson, L. and Deary, I. (2004) Intelligence predicts health and longevity, but why? *Curr. Dir. Psychol. Sci.*, **13**, 1-4.
- Greenland, S. (2000) Causal analysis in the health sciences. *J. Am. Statist. Ass.*, **95**, 286-289.
- Greenland, S. and Brumback, B. (2002) An overview of relations among causal modelling methods. *Int. J. Epidemiol.*, **31**, 1030-1037.
- Grossman, M. (1972) On the concept of health capital and the demand for health. *J. Polit. Econ.*, **80**, 223-255.
- Hart, C.L., Taylor, M.D., Davey Smith, G., Whalley, L.J., Starr, J.M., Hole, D.J., Wilson, V. and Deary, I.J. (2003) Childhood IQ, social class, deprivation, and their relationships with mortality and morbidity risk in later life: prospective observational study linking the Scottish Mental Survey 1932 and the Midspan studies. *Psychosom. Med.*, **65**, 877-883.

- Hawkes, D. and Plewis, I. (2004) Modelling attrition in the National Child Development Study. Presented at the Royal Statistical Society conference on Statistical Methods for Attrition and Non-response in Social Surveys, 28 May 2004 (<http://www.ccsr.ac.uk/methods/events/attrition/documents/1> - accessed 21/7/2004).
- Hotopf, M., Mayou, R., Wadsworth, M. and Wessely, S. (1999) Childhood risk factors for adults with medically unexplained symptoms: results from a national birth cohort study. *Am. J. Psychiatry*, **156**, 1796-1800.
- Hu, L. and Bentler, P.M. (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Modeling*, **6**, 1-55.
- Isohanni, I., Jobes, P.B., Jarvelin, M.R., Nieminen, P., Rantakallio, P., Jokelainen, J., Croudace, T.J. and Isohanni, M. (2001) Educational consequences of mental disorders treated in hospital. A 31-year follow-up of the Northern Finland 1966 Birth Cohort. *Psychol. Med.*, **31**, 339-349.
- Jefferis, B.J., Power, C. and Hertzman, C. (2002) Birth weight, childhood socioeconomic environment, and cognitive development in the 1958 British birth cohort study. *Brit. Med. J.*, **325**, 305.
- Keneckel, D. (1991) Health behaviour, health knowledge and schooling. *J. Polit. Econ.*, **99**, 287-305.
- Lewis, S., Ross, C.E. and Mirowsky, J. (1999) Establishing a sense of personal control in the transition to adulthood. *Soc. Forces*, **77**, 1573-1599.
- Little, R.J.A. and Rubin, D.B. (2002) *Statistical Analysis with Missing Data* (second edition). Hoboken, NJ: Wiley.
- Mirowsky, J. (1995) Age and the sense of control. *Soc. Psychol. Quart.*, **58**, 31-43.
- Mirowsky, J. and Ross, C.E. (1998) Education, personal control, lifestyle and health: a human capital hypothesis. *Res. Aging*, **20**, 415-449.
- Mirowsky, J. and Ross, C.E. (2003) *Social Status and Health*. New York: Aldine de Gruyter.
- Morris, J.N., Blane, D.B. and White, I.R. (1996) Levels of mortality, education, and social conditions in the 107 local authority areas of England. *J. Epidemiol. Commun. H.*, **50**, 15-17.

- Murasko, J. (2003) The human capital model of health: childhood origins. Working paper ([http://www.ruf.rice.edu/~muraskoj/html/murasko\\_paper1.pdf](http://www.ruf.rice.edu/~muraskoj/html/murasko_paper1.pdf) - accessed 21/7/04).
- Muthén, B.O. (2004) Mplus technical appendices (<http://www.statmodel.com/mplus/techappen.pdf> - accessed 17/10/04).
- Muthén, L.K. and Muthén, B.O. (2003) *Mplus User's Guide* (second edition). Los Angeles, CA: Muthén & Muthén.
- Pearl, J. (1995) Causal diagrams for empirical research. *Biometrika*, **82**, 669-710.
- Pearl, J. (2001) Direct and indirect effects. In: *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, 411-420. San Francisco, CA: Morgan Kaufman.
- Pearlin, L.I., Lieberman, M.A., Menaghan, E.G. and Mullan, J.T. (1981) The stress process. *J. Health. Soc. Behav.*, **22**, 337-356.
- Power, C., Manor, O., Fox, A.J., Fogelman, K. (1990) Health in childhood and social inequalities in health among young adults. *J. R. Statist. Soc. A*, **153** 17-28.
- Power, C. and Hertzman, C. (1997) Social and biological pathways linking early life and adult disease. *Brit. Med. Bull.*, **53**, 210-222.
- Ross, C.E. and Wu, C. (1995) The links between education and health. *Am. Sociol. Rev.*, **60**, 719-745.
- Rotter, J.B. (1966) Generalized expectancies for internal versus external control of reinforcement. *Psychol. Monogr.*, **80**, 1-28.
- Sandana, R., Mathers, C.M., Lopez, A.D., Murray, C.J.L. and Iburg, K.M. (2002) Comparative analyses of more than 50 household surveys on health status. In: Murray, C.J.L., Salomon, J.A., Mathers, C., Lopez, A.D. (eds.), *Summary Measures of Population Health, Concepts, Ethics, Measurement and Applications*. Geneva: World Health Organization.
- Seeman, M. and Seeman, T. (1983) Health behaviour and personal autonomy: a longitudinal study of the sense of control in illness. *J. Health Soc. Behav.*, **24**, 144-160.
- Singh-Manoux, A., Clarke, P. and Marmot, M.G. (2002) Multiple measures of socioeconomic position and psychosocial health: proximal and distal effects. *Int. J. Epidemiol.*, **31**, 1192-1199.

- Smith, G.D., Hart, C.L., Hole, D.J., MacKinnon, P., Gillis, C., Watt, G., Blane, D. and Hawthorne, V. (1998) Education and occupational social class: which is the more important indicator of mortality risk? *J. Epidemiol. Commun. H.*, **52**, 153-160.
- Sobel, M.E. (2000) Causal inference in the social sciences. *J. Am. Statist. Ass.*, **95**, 647-651.
- Wadsworth, M.E., Butterworth, S.L., Hardy, R.J., Kuh, D.J., Richards, M., Langenberg, C., Hilder, W.S. and Connor, M. (2003) The life course prospective design: an example of the benefits and problems associated with study longevity. *Soc. Sci. Med.*, **57**, 2193-2205.
- White, I., Blane, D. and Morris, J. (1999) Educational attainment, deprivation-affluence and self-reported health in Britain: a cross-sectional study. *J. Epidemiol. Commun. H.*, **53**, 535-541.
- Winkleby, M.A., Jatulis, D.E., Frank, E. and Fortmann, S.P. (1992) Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *Am. J. Public Health*, **57**, 153-182.

Table 1 Unstandardised probit regression coefficients and goodness of fit statistics of the latent variables Adult Health, Adolescent Health, Healthy Behaviours and Sense of Control.

	Women	Men
<b>(Good) Health at age 42 by</b>		
Good Health in general <sup>1</sup>	1.00	1.00
Good Health over the past year	0.96 (0.92-1.01)	0.99 ((0.94-1.04)
No Limiting long standing illness	0.72 (0.68-0.76)	0.66 (0.61-0.70)
No Hospital admissions (not accidents)	0.22 (0.18-0.26)	0.20 (0.16-0.25)
chisq/df	3.32, 2	13.86, 2
CFI	1.00	1.00
TLI	1.00	1.00
RMSEA	0.01	0.03
<b>(Good) Adolescent health at age 16 by</b>		
3 or Fewer GP Visits in last year <sup>1</sup>	1.00	1.00
Not Disabled or Handicapped	0.82 (0.54-1.10)	0.57 (0.41-0.73)
No Hospital Admission for 1 or more nights	0.98 (0.66-1.31)	0.86 (0.68-1.05)
Ill-health: No school absence over 1 month	1.74 (1.08-2.39)	1.27 (0.97-1.57)
chisq/df	4.60, 2	1.61, 2
CFI	0.99	1.00
TLI	0.96	1.00
RMSEA	0.01	0.00
<b>Healthy behaviours at age 42 by</b>		
Some Regular Exercise <sup>1</sup>	1.00	1.00
Fruit consumption more than 1/week	1.26 (1.04-1.48)	1.68 (1.36-2.00)
Not Smoking	1.22 (1.01-1.43)	1.40 (1.15-1.65)
chisq/df	0,0	0,0
CFI	1.00	1.00
TLI	1.00	1.00
RMSEA	0.00	0.00
<b>(Lack of) Sense of control at age 33 by</b>		
Usually have no choice&control over life <sup>1</sup>	1.00	1.00
Never get what I want out of life	1.26 (1.18-1.35)	1.29 (1.16-1.41)
Find life's problems too much for me	0.75 (0.70-0.80)	0.55 (0.50-0.60)
chisq/df	0,0	0,0
CFI	1.00	1.00
TLI	1.00	1.00
RMSEA	0.00	0.00

\*\* p <0.01

<sup>1</sup>Reference indicator for identifying latent variable

chisq/df: chisquare/degrees of freedom

CFI: Comparative Fit Index

TLI: Tucker Lewis index

RMSEA: Root Mean Square Error of Approximation

Table 2. Unstandardised regression coefficients for variables (Education, Healthy Behaviours, Cognitive Ability, Childhood Social Class, Adolescent Health, Sense of Control and Adult Social Class) on Adult Health. For men.

Variable	Bivariate Analysis		Adjusted Bivariate Analysis		
		Variable +	Variable +	Variable +	Variable +
		Cognitive Ability	Childhood Social Class	Adolescent Health	Sense of Control
Education	<b>0.34 (0.31, 0.38)</b>	<b>0.14 (0.09, 0.18)</b>	<b>0.17 (0.14, 0.20)</b>	<b>0.08 (0.04, 0.13)</b>	<b>0.11 (0.08, 0.14)</b>
Education-squared	<b>-0.01 (-0.01, -0.01)</b>	<b>-0.00 (-0.00, -0.00)</b>	<b>-0.00 (-0.01, -0.00)</b>	0.00 (0.00, 0.00)	<b>-0.00 (-0.01, -0.00)</b>
Cognitive Reading Score	<b>3.54 (1.96, 5.12)</b>	<b>0.03 (0.01, 0.05)</b>			
Cognitive Mathematics Score	<b>0.18 (0.16, 0.20)</b>	<b>0.01 (0.00, 0.02)</b>			
Childhood Social Class	<b>-0.15 (-0.18, -0.12)</b>		-0.03 (-0.06, 0.00)		
Adolescent Health	<b>0.59 (0.42, 0.76)</b>			<b>0.98 (0.59, 1.38)</b>	
Sense of Control (low)	<b>-0.73 (-0.81, -0.65)</b>				<b>-0.36 (-0.44, -0.29)</b>
Adult Social Class	<b>-0.42 (-0.46, -0.37)</b>				
Healthy Behaviours	<b>1.14 (0.99, 1.29)</b>				

Variable	Adjusted Bivariate analysis		Direct Effects	Causal Analysis	
	Variable +	Variable +		Indirect Effects	Total Effects
	Adult Social Class	Healthy Behaviours			
Education	<b>0.14 (0.10, 0.17)</b>	<b>0.05 (0.01, 0.09)</b>	-0.04 (-0.11, 0.04)	<b>0.14 (0.10, 0.18)</b>	<b>0.10 (0.04, 0.17)</b>
Education-squared	<b>-0.00 (-0.01, -0.00)</b>	<b>-0.00 (-0.00, -0.00)</b>	0.00 (-0.01, 0.00)	<b>-0.00 (-0.00, -0.00)</b>	<b>-0.00 (-0.01, -0.00)</b>
Cognitive Reading Score			-0.01 (-0.03, 0.01)	<b>0.02 (0.01, 0.03)</b>	0.01 (-0.01, 0.03)
Cognitive Mathematics Score			0.00 (-0.01, 0.01)	0.00 (0.00, 0.01)	0.00 (0.00, 0.01)
Childhood Social Class			0.03 (-0.01, 0.06)	<b>-0.04 (-0.06, -0.02)</b>	-0.02 (-0.05, 0.02)
Adolescent Health			<b>1.28 (0.58, 1.99)</b>	<b>0.48 (0.20, 0.75)</b>	<b>1.76 (0.98, 2.54)</b>
Sense of Control (low)			<b>-0.25 (-0.35, -0.16)</b>	<b>-0.06 (-0.09, -0.02)</b>	<b>-0.31 (-0.40, -0.22)</b>
Adult Social Class	<b>-0.10 (-0.13, -0.06)</b>		<b>-0.04 (-0.08, 0.00)</b>	<b>-0.02 (-0.04, -0.01)</b>	<b>-0.06 (-0.10, -0.02)</b>
Healthy Behaviours		<b>0.68 (0.54, 0.83)</b>	<b>0.48 (0.31, 0.66)</b>	0.00	<b>0.48 (0.31, 0.66)</b>

Notes: 1. Bivariate analysis: All models of the form Adult Health = Variable, except Adult Health = Education + Education-squared. 2. Adjusted Bivariate analysis: All models of the form Adult Health = Variable + X, where X = Cognitive Ability, Childhood Social Class, Adolescent Health, Sense of Control, Adult Social Class, and Healthy Behaviours, except Adult Health = Education + Education-squared + X. 3. Estimated direct effect of variable on Health in full SEM. See Figure 2 for parameter estimates. 4. 95 percent confidence intervals in parentheses, bold figures indicate statistical significance at 5 percent level.

Table 3. Unstandardised regression coefficients for variables (Education, Healthy Behaviours, Cognitive Ability, Childhood Social Class, Adolescent Health, Sense of Control and Adult Social Class) on Adult Health. For women.

Variable	Bivariate Analysis		Adjusted Bivariate Analysis		
		Variable + Cognitive Ability	Variable + Childhood Social Class	Variable + Adolescent Health	Variable + Sense of Control
Education	<b>0.20 (0.17, 0.23)</b>	<b>0.15 (0.09, 0.21)</b>	<b>0.15 (0.19, 0.22)</b>	<b>0.11 (0.07, 0.15)</b>	<b>0.15 (0.12, 0.19)</b>
Education-squared	<b>-0.01 (-0.01, -0.00)</b>	<b>-0.00 (-0.01, -0.00)</b>	<b>-0.01 (-0.01, -0.00)</b>	<b>-0.00 (-0.01, -0.00)</b>	<b>-0.01 (-0.01, -0.00)</b>
Cognitive Reading Score	<b>0.06 (0.04, 0.07)</b>	<b>0.03 (0.00, 0.05)</b>			
Cognitive Mathematics Score	<b>0.03 (0.02, 0.04)</b>	<b>0.01 (0.00, 0.02)</b>			
Childhood Social Class	<b>-0.13 (-0.15, -0.10)</b>		<b>-0.08 (-0.05, -0.01)</b>		
Adolescent Health	<b>0.62 (0.48, 0.75)</b>			<b>0.59 (0.42, 0.76)</b>	
Sense of Control	<b>-0.36 (-0.42, -0.30)</b>				<b>-0.31 (-0.38, -0.24)</b>
Adult Social Class	<b>-0.10 (-0.13, -0.08)</b>				
Healthy Behaviours	<b>0.98 (0.81, 1.15)</b>				

Variable	Adjusted Bivariate Analysis		Direct Effects	Causal Analysis Indirect Effects	Total Effects
	Variable + Adult Social Class	Variable + Healthy Behaviours			
Education	<b>0.19 (0.16, 0.23)</b>	-0.05 (0.01, 0.06)	<b>-0.12 (-0.20, -0.03)</b>	<b>0.18 (0.13, 0.24)</b>	<b>0.07 (0.01, 0.13)</b>
Education-squared	<b>-0.01 (-0.01, -0.00)</b>	0.00 (0.00, 0.00)	0.00 (0.00, 0.01)	<b>-0.00 (-0.01, -0.00)</b>	0.00 (-0.01, 0.00)
Cognitive Reading Score			0.01 (-0.01, 0.03)	0.01 (-0.01, 0.02)	0.02 (0.00, 0.03)
Cognitive Mathematics Score			<b>0.01 (0.00, 0.02)</b>	0.00 (-0.01, 0.00)	<b>0.01 (0.00, 0.02)</b>
Childhood Social Class			-0.01 (-0.05, 0.02)	<b>-0.05 (-0.07, -0.03)</b>	<b>-0.06 (-0.10, -0.03)</b>
Adolescent Health			<b>0.65 (0.43, 0.87)</b>	<b>0.13 (0.03, 0.24)</b>	<b>0.78 (0.58, 0.98)</b>
Sense of Control			<b>-0.14 (-0.22, -0.05)</b>	<b>-0.14 (-0.19, -0.09)</b>	<b>-0.27 (-0.34, -0.20)</b>
Adult Social Class	-0.06 (-0.03, 0.01)		0.01 (-0.03, 0.04)	-0.01 (-0.03, 0.01)	0.00 (-0.04, 0.03)
Healthy Behaviours		<b>1.03 (0.80, 1.26)</b>	<b>0.95 (0.68, 1.22)</b>	0.00	<b>0.95 (0.68, 1.22)</b>

Notes: 1. Bivariate analysis: All models of the form Adult Health = Variable, except Adult Health = Education + Education-squared. 2. Adjusted Bivariate analysis: All models of the form Adult Health = Variable + X, where X = Cognitive Ability, Childhood Social Class, Adolescent Health, Sense of Control, Adult Social Class, and Healthy Behaviours, except Adult Health = Education + Education-squared + X. 3. Estimated direct effect of variable on Health in full SEM. See Figure 3 for parameter estimates. 4. 95 percent confidence intervals in parentheses, bold figures indicate statistical significance at 5 percent level.

Appendix 1a

Distribution of all the variables in the analysis

	<b>Region</b>		<b>Low Birth Weight</b>		<b>Gender</b>
	<i>Age 0</i>		<i>Age 0</i>		<i>Age 0</i>
England	14517	lowest quintile	13931	Male	8999
Wales	914	normal	3307	Female	8412
Scotland	1985	Missing	178	Missing	5
Total	17416	Total	17416	Total	17416

	<b>Cognition- Reading score</b>		<b>Father's Social Class</b>		<b>Cognition- Math Score</b>
	<i>Age 7</i>		<i>Age 7</i>		<i>Age 7</i>
0	64	Professional	923	0	348
1	9	Managerial	2419	1	653
2	16	Skilled NM	1698	2	1332
3	35	Skilled manual	8060	3	1756
4	51	Partly skilled	2905	4	2017
5	85	Unskilled	1326	5	1979
6	131	Missing	85	6	1910
7	163	Total	17416	7	1624
8	184			8	1278
9	215			9	911
10	188			10	576
11	243			Missing	3032
12	265			Total	17416
13	222				
14	257				
15	296				
16	273				
17	289				
18	340				
19	363				
20	401				
21	393				
22	458				
23	475				
24	586				
25	674				
26	802				
27	1000				
28	1298				
29	1923				
30	2718				
Missing	2999				
Total	17416				

		<b>Age Mum Left</b>	
		<b>Full Time Education</b>	<i>Age 16</i>
		under 13 yrs	65
		13 to 14 years	91
		14 to 15 years	5078
		15 to 16 years	3291
		16 to 17 years	1232
		17 to 18 years	453
		18 to 19 years	243
		19 to 21 years	137
		21 to 23 years	167
		23 or more years	42
		Missing	6617
		Total	17416

Appendix 1b Distribution of Variables in the Analysis

	<b>GP visits</b> <i>Age 16</i>	<b>Disability/ Handicap</b> <i>Age 16</i>	<b>Hospital Admissions</b> <i>Age 16</i>	<b>Illness school Absence</b> <i>Age 16</i>
Poor health	8962	9487	9877	9678
Good health	1304	786	623	1042
Missing	7150	7143	6916	6696
Total	17416	17416	17416	17416

	<b>Education Qualifications</b> <i>Age 23</i>		<b>Social Class</b> <i>Age 33</i>
No qualification	1436	Professional	498
CSE 2-5/equiv NVQ1	1382	Managerial	2717
O Level/equiv NVQ2	3660	Skilled NM	2641
A Level/equiv NVQ3	1748	Skilled manual	1929
Higher qual NVQ4	1000	Partly skilled	1966
Degree or higher NVQ5,6	911	Unskilled	558
Missing	7279	Missing	7107
Total	17416	Total	17416

	<b>Free choice &amp; control over life</b> <i>Age 33</i>	<b>Get what I want out of life</b> <i>Age 33</i>	<b>Run my life as I want to</b> <i>Age 33</i>
Yes	9313	Yes	9856
No	1255	No	704
Missing	7992	Missing	8000
Total	18560	Total	18560

	<b>Self-rated Health</b> <i>Age 41</i>	<b>Health over last 12 months</b> <i>Age 41</i>	<b>Limiting Health</b> <i>Age 41</i>	<b>Hospital Admissions</b> <i>Age 41</i>
Poor health	1948	2776	9494	3637
Good health	8838	8011	1284	7145
Missing	6630	6629	6638	6634
Total	17416	17416	17416	17416

	<b>Smoking status</b> <i>Age 41</i>	<b>Eat Fresh Fruit</b> <i>Age 41</i>	<b>Exercise</b> <i>Age 41</i>
non-smoker	7993	once a week	8002
current smoker	2793	lt once a week	2781
Missing	6630	Missing	6633
Total	17416	Total	17416

Figure 1a: All Associations Between Education and Health

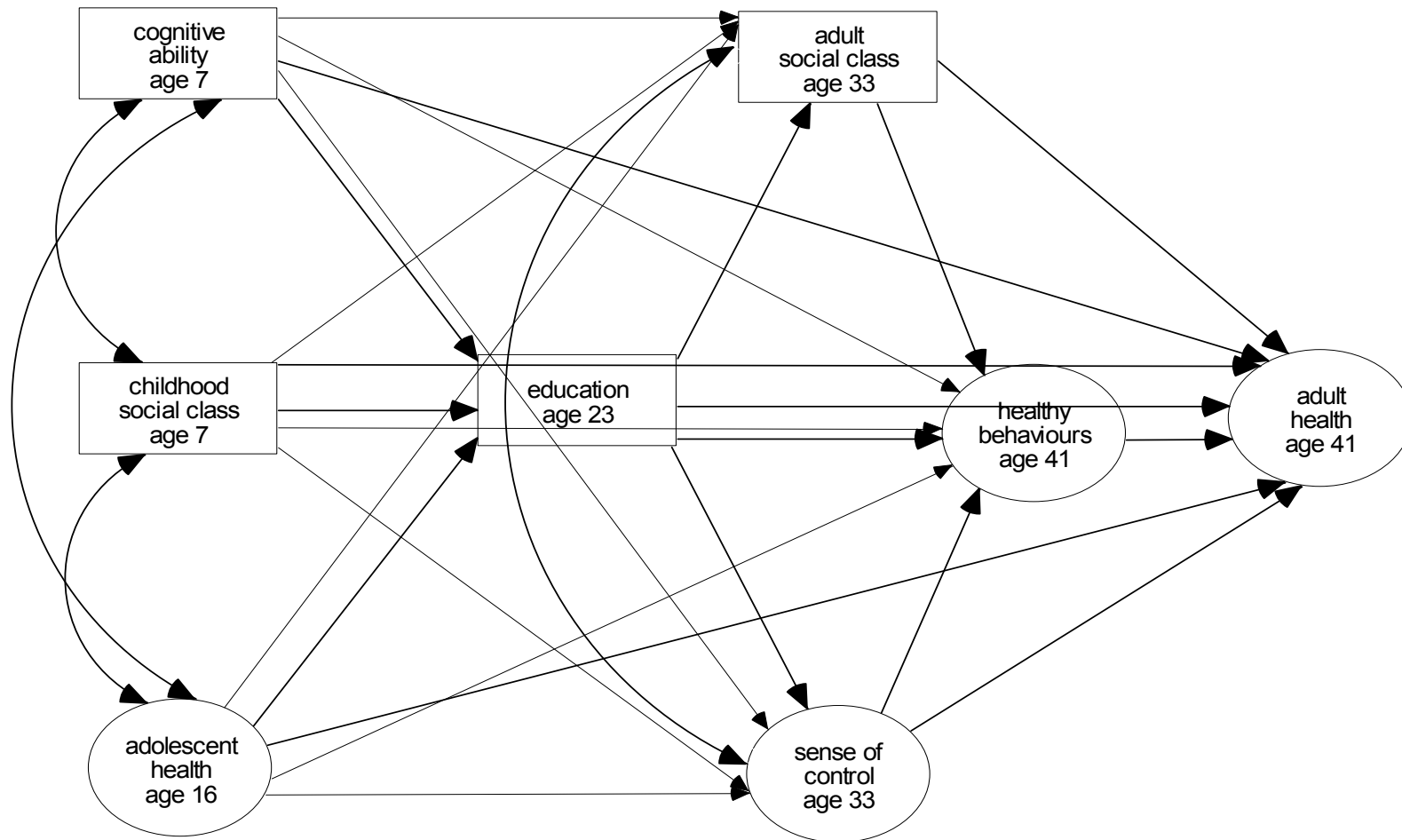


Figure 1b: Principle Causal Mechanisms underlying Education and Health

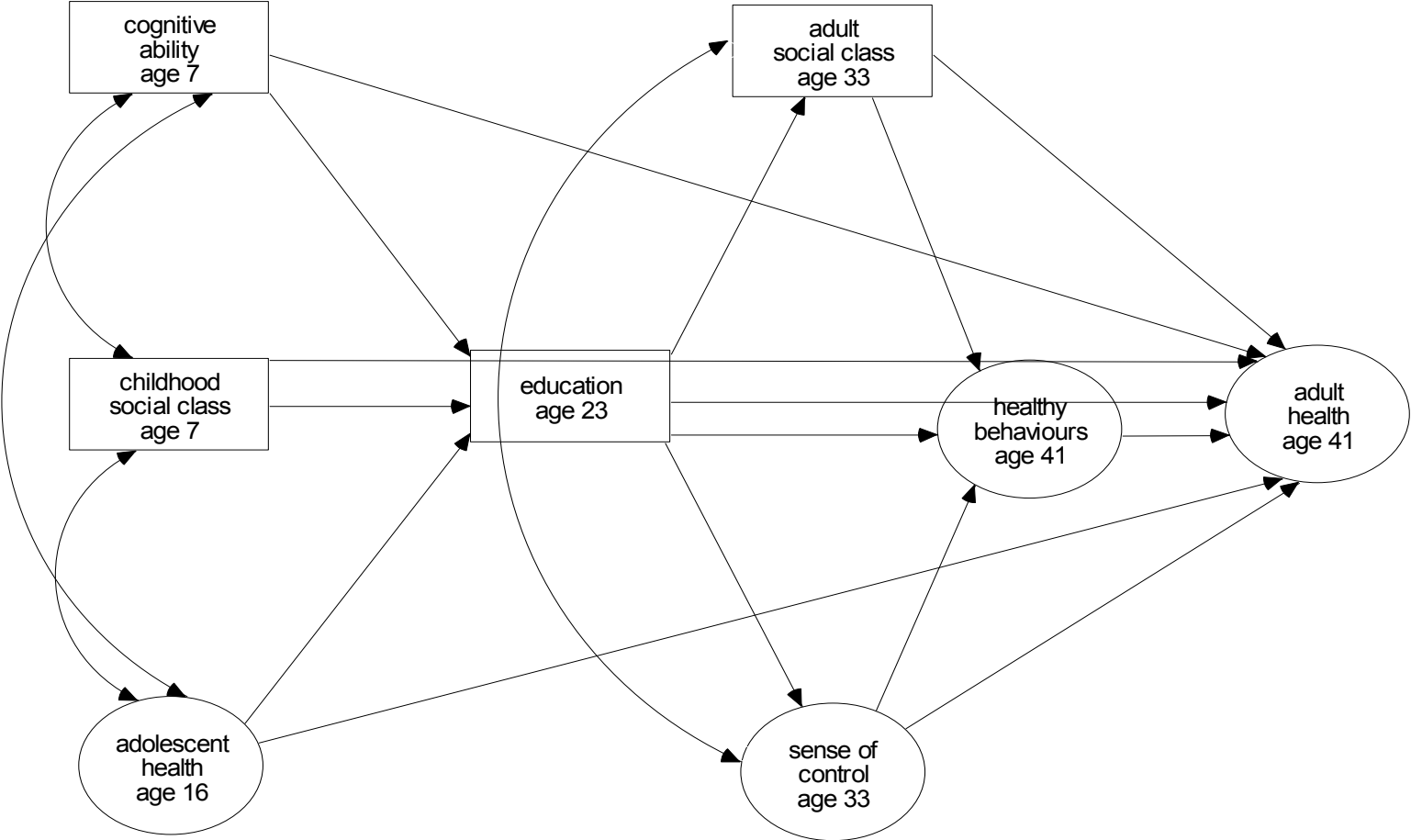
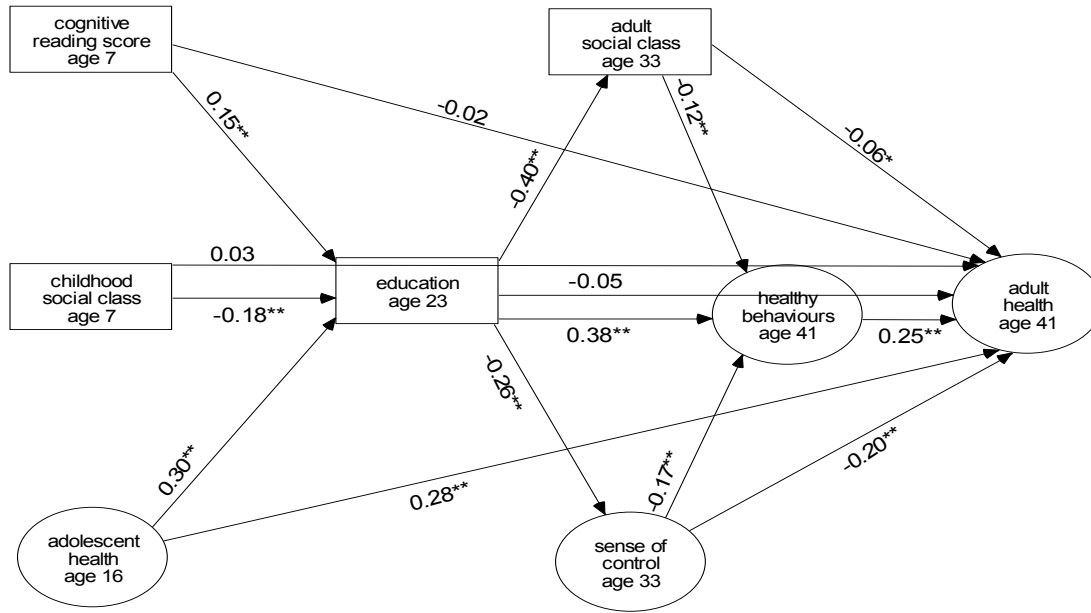
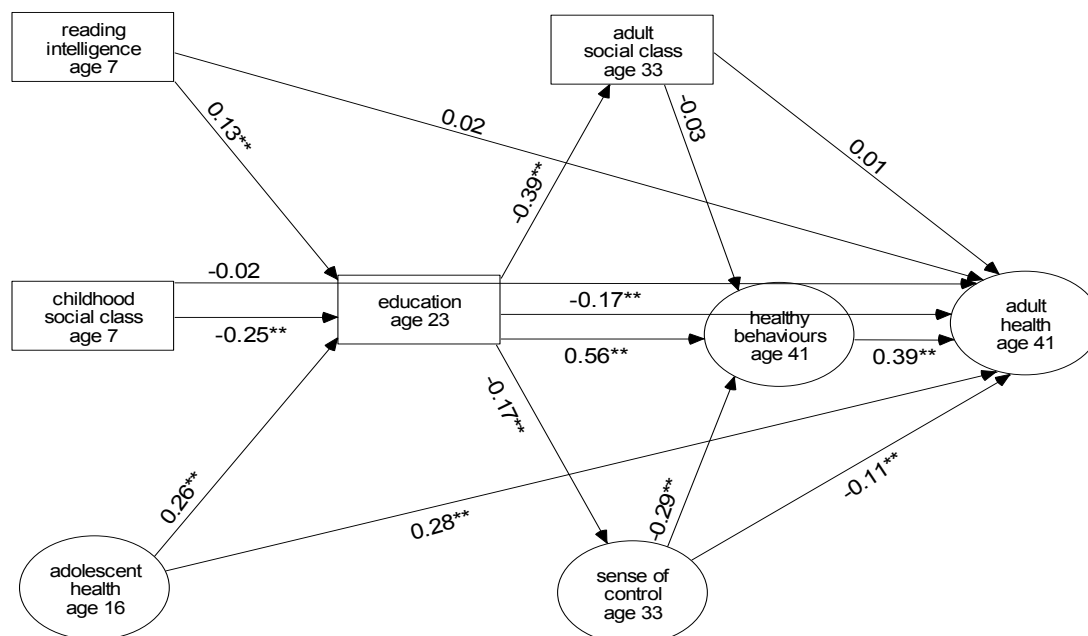


Figure 2 Standardised Regression Coefficients (and Correlations) of the Mechanisms Linking Education and Health- Men



	Low B Wt age 0	Wales age 0	Scotland age 0	Social cls age 7	Read age 7	Math age 7	Mum's ed. age 16	Health age 16	Quals age 23	Quals sq age 23	Social cls age 33	Control age 33	Healthy age 42	Health age 42
Low Birth Weight- age 0	1.00													
Born in Wales- age 0	-0.07	1.00												
Born in Scotland- age 0	0.01	<b>-0.58</b>	1.00											
Childhood class- age 7	<b>0.08</b>	0.01	<b>0.11</b>	1.00										
Ability Reading- age 7	<b>-0.13</b>	<b>0.08</b>	<b>-0.05</b>	<b>-0.21</b>	1.00									
Ability Math- age 7	<b>-0.13</b>	-0.04	<b>0.17</b>	<b>-0.26</b>	<b>0.56</b>	1.00								
Mother's education- age 16	<b>-0.09</b>	<b>0.07</b>	<b>-0.12</b>	<b>-0.37</b>	<b>0.20</b>	<b>0.21</b>	1.00							
Child health- age 16	<b>-0.16</b>	<b>-0.13</b>	0.01	<b>-0.24</b>	<b>0.25</b>	<b>0.33</b>	<b>0.16</b>	1.00						
Quals- age 23	0.01	-0.02	0.01	<b>-0.18</b>	<b>0.15</b>	<b>0.23</b>	<b>0.09</b>	<b>0.30</b>	1.00					
Quals sq- age 23	<b>0.14</b>	-0.04	<b>0.12</b>	<b>0.05</b>	0.03	<b>-0.11</b>	0.04	0.00	<b>0.36</b>	1.00				
Social class- age 33	0.03	0.03	<b>0.05</b>	<b>0.13</b>	-0.02	<b>-0.09</b>	<b>-0.04</b>	0.01	<b>-0.40</b>	0.01	1.00			
Control- age 33	-0.04	-0.05	<b>-0.08</b>	0.01	<b>-0.13</b>	0.02	0.00	-0.10	<b>-0.26</b>	0.03	<b>0.08</b>	1.00		
Healthy behavrs- age 42	-0.03	0.06	0.00	-0.04	-0.02	<b>-0.07</b>	-0.04	<b>0.15</b>	<b>0.38</b>	<b>-0.17</b>	<b>-0.12</b>	<b>-0.17</b>	1.00	
Adult health- age 42	0.03	0.03	-0.01	0.03	-0.02	0.00	0.02	<b>0.28</b>	-0.05	-0.06	<b>-0.06</b>	<b>-0.20</b>	<b>0.25</b>	1.00

Figure 3 Standardised Regression Coefficients (and Correlations) of the Mechanisms Linking Education and Health- Women



	Low B Wt age 0	Wales age 0	Scotland age 0	Social cls age 7	Read age 7	Math age 7	Mum's ed. age 16	Health age 16	Quals age 23	Quals sq age 23	Social cls age 33	Control age 33	Healthy age 42	Health age 42
Low Birth Weight- age 0	1.00													
Born in Wales- age 0	0.00	1.00												
Born in Scotland- age 0	-0.01	<b>-0.58</b>	1.00											
Childhood class- age 7	<b>0.07</b>	<b>0.06</b>	<b>0.10</b>	1.00										
Ability Reading- age 7	<b>-0.14</b>	0.05	<b>-0.07</b>	<b>-0.21</b>	1.00									
Ability Math- age 7	<b>-0.16</b>	-0.04	<b>0.13</b>	<b>-0.27</b>	<b>0.53</b>	1.00								
Mother's education- age 16	<b>-0.04</b>	0.02	<b>-0.07</b>	<b>-0.37</b>	<b>0.15</b>	<b>0.19</b>	1.00							
Child health- age 16	-0.02	-0.08	-0.06	<b>-0.17</b>	<b>0.17</b>	<b>0.25</b>	<b>0.14</b>	1.00						
Quals- age 23	<b>-0.07</b>	0.02	<b>0.06</b>	<b>-0.25</b>	<b>0.13</b>	<b>0.25</b>	<b>0.15</b>	<b>0.26</b>	1.00					
Quals sq- age 23	<b>0.09</b>	<b>-0.10</b>	<b>0.08</b>	-0.03	<b>-0.05</b>	<b>-0.04</b>	<b>0.05</b>	0.00	<b>0.42</b>	1.00				
Social class- age 33	0.03	0.01	-0.01	<b>0.07</b>	-0.03	<b>-0.07</b>	-0.01	0.01	<b>-0.39</b>	0.00	1.00			
Control- age 33	0.01	-0.04	0.03	0.04	<b>-0.09</b>	-0.02	-0.03	<b>-0.12</b>	<b>-0.17</b>	-0.06	<b>0.08</b>	1.00		
Healthy behavrs- age 42	-0.01	-0.04	<b>-0.13</b>	-0.08	-0.05	<b>-0.10</b>	-0.03	0.02	<b>0.56</b>	<b>-0.31</b>	-0.03	<b>-0.29</b>	1.00	
Adult health- age 42	-0.02	0.04	0.01	-0.02	0.02	<b>0.07</b>	0.02	<b>0.28</b>	<b>-0.17</b>	0.04	<b>0.01</b>	<b>-0.11</b>	<b>0.39</b>	1.00