I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12439

Widening the High School Curriculum to Include Soft Skill Training: Impacts on Health, Behaviour, Emotional Wellbeing and Occupational Aspirations

Grace Lordan
Alistair McGuire

JUNE 2019

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12439

Widening the High School Curriculum to Include Soft Skill Training: Impacts on Health, Behaviour, Emotional Wellbeing and Occupational Aspirations

Grace Lordan
LSE and IZA

Alistair McGuire
LSE

JUNE 2019

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Widening the High School Curriculum to Include Soft Skill Training: Impacts on Health, Behaviour, Emotional Wellbeing and Occupational Aspirations*

From 2020 Personal, Social, Health and Economic Education will be compulsory in UK schools for adolescents, however less is known about how it can be taught in a an effective manner. We examine, through a randomised trial, the impact of an evidenced based health related quality of life (HRQoL) curriculum called *Healthy Minds* that ran in 34 high schools in England over a four-year period. We find robust evidence that *Healthy Minds* positively augments many physical health domains of treated adolescents. We also find some evidence that *Healthy Minds* positively affects behaviour, but has no impact on emotional wellbeing. We find notable gender effects, strongly favouring boys. We also present evidence that *Healthy Minds* changes career aspirations, with those exposed to treatment being less likely to choose competitive work and more likely to choose work that involves "*people*-skills". Overall our work illustrates the potential for later childhood interventions to promote HRQoL and develop the career aspirations of adolescents.

JEL Classification: I18, I20

Keywords: soft skills, health related quality of life, character, high school curriculum, personal, social, health and economic education

Corresponding author:
Grace Lordan
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
United Kingdom

E-mail: g.lordan@lse.ac.uk

* We are very grateful for detailed comments received from Richard Layard, Richard Blundell, Steve Machin, Sandra McNally, Carol Propper, Guillermo Rodriguez-Guzman and Ian Walker as well as participants at the LSE Department of Health Policy, London Health Economics Group, and anonymous referees used by the EEF. A special thanks to Bounce Forward for delivering the program, HcareSolutions for independently collecting the data and Amy Challen for her invaluable research assistance at the initial phase of this project. We acknowledge funding from the UK Education Endowment Fund (EEF), Templeton Trust and the LSE.

1. Introduction and Background

From 2020 Personal, Social, Health and Economic Education will be compulsory in UK schools for adolescents, however less is known about how it can be taught in an effective manner. In essence this type of education aims to augment Health related quality of life (HRQoL). Specifically, HRQoL in childhood is a multi-dimensional concept including domains related to physical, mental, emotional, and social functioning that has a long and influential arm into adulthood[1]. Recently it has also been noted that soft skills, a component of HRQoL, are increasingly in demand in the labour market given these skills are not readily substitutable with technology (Autor and Dorn, 2013; Autor and Dorn, 2015), and are particularly valuable for children leaving school with a high school diploma or less.[2] There is also evidence that other dimensions of HRQoL, such as poor adolescent mental health, correlate with long-term negative impacts into adulthood, with impacts on educational attainment (Gibb *et al.*, 2012), the ability to work and earn as an adult (Goodman et al, 2011) and long run psychological disturbance (Collishaw et al, 2004; Thapar et al, 2012).

Given the importance of HRQoL a natural question arises as to whether this specific form of human capital can be augmented throughout childhood within the public education system. Schools provide a major opportunity for this type of investment, and elements of HRQoL, specifically those related to soft skills, (as compared to cognitive skills) have been shown to be malleable throughout the life course (Heckman and Kautz (2013). Proving any programme's effectiveness is key given that there is as yet no generally accepted way to teach HRQoL in schools, which in itself is a barrier to having HRQoL programs on any teaching curriculum. While such programmes could crowd out traditional academic achievement, we know of no analysis that forwards evidence (compelling or otherwise) that this is the case. In fact, Heckman and Kutz (2014) and Kautz, Heckman, Diris et al (2014) provide general

[1] This conclusion reflects the findings of many influential papers, which highlight that childhood health significantly predicts adult labour market outcomes (Case *et al.*, 2005; Black *et al.*, 2007; Smith, 2009; Currie *et al.*, 2009; Currie, 2009, Case and Paxson, 2011; Case and Paxson, 2010 and Lekfuangfu and Lordan, 2018).

[2] Lordan and Neumark (2018) highlight that jobs at the low end of the distribution are disappearing quicker, in the face of rising labor costs, if they can be automated by robotics. The low skilled occupations that remain and are growing involve interactions with people. In contract, high skilled jobs that remain either involve people and/or abstract thinking. This gives more choice for individual specialization. This conclusion is supported by a number of other studies (see Acemoglu, 2002; Autor, Levy and Murnane, 2003; Beckman, Bound and Machin, 1998)

literature reviews where they outline mechanisms and empirical support to suggest that augmenting soft skills through a school-based intervention actually improves (rather than crowds out) educational achievement.

We are particularly interested in whether HRQoL can be changed within the secondary (high) school setting through targeting adolescents, given that we know of no study to offer any evidence in this regard. The importance of this question is echoed in the fact that we are evaluating these children more than halfway through their childhood, at a point when it may be difficult to amplify cognitive skills, and where to date it remains unknown whether a programme specifically designed to augment HRQoL can allow children to catch up across its many domains. To this end we analyse the effectiveness of a four-year, evidenced-based programme administered within schools in England, referred to as the *Healthy Minds* programme, designed specifically to alter the HRQoL trajectory of children aged 11-12 years.

There is a growing literature that considers interventions that roll out in earlier childhood, but none that consider interventions aimed at adolescents in a school setting. Our work contributes to this evidence base. Notably, the majority of work on childhood interventions is U.S. based and targets pre-adolescent children. This evidence base suggests that such programmes can succeed in their goals (see Durlak, Weissberg, Dymnicki, Taylor, and Schellinger, (2011), for a meta-analysis of 213 school-based social and emotional learning programmes), with studies encompassing a longer follow up having a mean impact on various aspects of HRQoL that are positive and statistically significant.[3] Directly related to our work is a strand of research by Cuhna and Heckman (2008) and Cuhna, Heckman and Schennach, (2010), who develop a household production conceptual approach, to provide a framework within which to evaluate cognitive and non-cognitive childhood investments on longer term outcomes. This framework has been used by Almond, Currie and Duque (2018) to extensively review the literature in this area, with a specific focus on pre-school and early childhood interventions. Their review highlights that such interventions generally show positive impact on personality traits,

[3] See for example evidence for the long run impacts of the Perry Preschool program in Heckman, Pinto, and Savelyev (2013) and The Abecedarian Program in Campbell, Conti, Heckman, Moon, and Pinto (2013)..

health behaviours and health outcomes (see for specific examples Heckman, Pinto and Savelyev, 2013; Campbell et al, 2014, Conti, Heckman and Pinto, 2016).

They caution, however, that outcomes can be detrimental if, for example, pre-school programmes are of lower quality to individuals than any alternative provision. In support of this particular conclusion they cite the work by Havnes and Mogstad (2015) who find pre-schooling increased the earning capacity of the poorest children but reduced the earnings capacity of the richest, and Baker, Gruber and Milligan (2015) who find that a universal child care programme had negative effects on children's noncognitive outcomes and significant declines in behaviour in boys and self-reported health generally. Overall however, the review by Almond, Currie and Duque (2018) of the early childhood intervention literature is one of positive effect.

In a separate review Heckman and Kautz (2013), support the general conclusion reached by Almond, Currie and Duque (2018) that well-designed childhood interventions can positively impact on health behaviours and outcomes. They provide a compelling argument backed by empirical evidence, that aspects of HRQoL, defined widely to incorporate aspects of behaviour and self-esteem, are skills rather than fixed traits that can be changed throughout the life course. Notably, the authors suggest that these are more malleable than cognitive skills in later childhood.

Our work also makes a direct contribution to public policy in the UK where the government has recently acknowledged that personal, social, health and economic education (PSHE) at school is a means to provide young people with the skills to become more self-aware, resilient and suitable in making more informed life-choices (House of Commons, 2015). This UK House of Commons Education Committee report also highlighted that the quality of PSHE, as currently taught in schools, is sub-optimal and teaching delivery in this area requires improvement in 40% of secondary schools (Ofsted, 2013). One of the recommendations of this Education Committee report was for PSHE, which during the period of our study was only a recommended subject area, to become a statutory, non-examined subject for all public sector high

school children, aged 11-16..[4] This has subsequently become a UK government objective. However, no evidence was provided on the content to be provided within any statutory curriculum.

The specific aim of this study is to evaluate a curriculum explicitly designed for the PSHE curriculum; the *Healthy Minds* programme. The *Healthy Minds* curriculum is an amalgam of 14 separate educational modules targeted at improving HRQoL, which include elements on building resilience, navigating social media, looking after mental health, developing healthy relationships and understanding the responsibilities of being a parent. The modules were selected if they had clear, demonstrable supportable evidence of their individual modular impact and met latest best practice guidance for delivering personal, social and health topics. Coleman et al. (2011) reviewed these individual modules and provides detail on how they were selected to be combined into the *Healthy Minds* curriculum. The curriculum was expressly designed to enhance the HRQoL of children in secondary (high) schools in the UK. Teachers delivering the programme were specifically trained in how to administer the teaching to ensure a level of consistent delivery; particularly important as PSHE is routinely taught by teachers with no specific training in the subject.

The *Healthy Minds* programme was introduced into English secondary (high) schools on an experimental basis and evaluated through conducting a randomised controlled trial (RCT). The experiment consisted of school cohorts characterised, in the main, by above average proportions of disadvantaged pupils. The RCT was conducted across 39 school-cohorts, and introduced in 2013 or 2014 during first year entry to state-funded secondary (high) school.[5] The *Healthy Minds* curriculum then ran for four years. The rollout of *Healthy Minds* was randomised to 23 school-cohorts, who followed the programme, while 16 school-cohorts did not receive the new curriculum and continued with the non-standardised PSHE offerings, acting as control cohorts for

[4] House of Commons Education Committee. Life lessons: PSHE and SRE in schools. Fifth Report of Session 2014-15. HC 145. London: HMSO

[5] The English educational system is compulsory up to the age 18, with schooling compulsory until 16, after which various academic or vocational education follows as determined by individual abilities. The secondary school system educates individuals from 11/12 years (school year 7) up to age 15/16, following primary school. At 15/16 pupils typically take the General Certificate of Secondary Education (GCSE) exam. *Healthy Minds* was introduced to pupils entering secondary schools and completed after four years, when pupils sat GCSE exams.

the experiment. The *Healthy Minds* curriculum was delivered in the treatment cohorts through weekly lessons of one hour each, which schools may have already allocated to PSHE education in their curriculum, although there was no obligation on schools to teach PHSE during the years through which our study was conducted.

While RCTs are a valuable approach, we recognise that they have limitations, including imperfections affecting random assignment, compliance, attrition and multiple hypothesis testing, especially when employed over a lengthy, four-year window. Consequently, as described fully below within the RCT framework, we evaluate findings on a (conservative) intent-to-treat basis, having used a claw-back mechanism to minimise attrition. We also consider a number of approaches to testing the robustness of this approach.

Precise measurement of HRQoL, our outcome of interest, is not straightforward. There is no single measured outcome that can capture HRQoL, so we rely on 13 sub-scales from a questionnaire validated in various childhood and adolescent studies, the CHQ-CF87 questionnaire (CHQ, 2013). This questionnaire captures health, behaviour and emotional wellbeing, as well as specific aspects relating to the child's family life. One advantage of this questionnaire is that it incorporates a general self-reported health measure that can be used to calculate the required sample size to detect a predetermined average treatment effect from the *Healthy Minds* programme. It also offers the advantage of allowing testing across a range of HRQoL dimensions through the use of various domain indicators. We adjust for multiple hypothesis testing in doing so.

Our empirical evidence provides strong support for the introduction of the *Healthy Minds* curriculum and is associated with a positive, statistically significant gain in our primary outcome measure of self-assessed health. Specifically, pupils exposed to *Healthy Minds* have global health attainment that is 0.235 standard deviations higher than children in the control group. We also find positive and significant gains to a number of other HRQoL domains. These are: physical difficulties, pain and discomfort, family activities and family cohesion scales to the order of 0.288, 0.239, 0.174 and 0.242 respectively. Finally, a summary of our results presents some evidence (significant at the 10% level of significance) that *Healthy Minds* positively

affects behaviour, yet overall has no impact on emotional wellbeing. A separate set of analyses illustrates that there are heterogenous treatment effects by gender, with boys having many more gains in HRQoL domains as compared to girls.

In an effort to examine the effects of *Healthy Minds* on labour market outcomes, in the absence of revealed preference data, we also study reported aspirations. Notably, our findings suggest that treatment with *Healthy Minds* negatively predicts the likelihood of choosing competitive work, but positively predicts the likelihood of choosing work that is high on '*people*' content. This seems to suggest that being exposed to the program moves pupils from a tendency to sort into jobs that are individualistic towards jobs that are more intrinsically motivated and, arguably, that will be of increasing importance over their lifespan.

2. The *Healthy Minds* Intervention

The *Healthy Minds* curriculum was a teaching programme specifically designed to deliver effective PSHE education in one-hour, weekly teaching slots within the normal secondary (high) school timetable over a four-year period. Secondary (high) school PSHE education in England is meant to provide a curriculum that teaches skills that lead to healthy development, either through raising awareness and prevention of detrimental health-related outcomes or the promotion of life skills aimed at encouraging healthy adolescent behaviour (MacDonald, 2008). However, while there is an established framework for PSHE education, was not statutory over the timespan of our study, nor taught within an agreed curriculum or to a universally high standard. Schools have generally given emphasis to STEM and other examined subjects, squeezing their teaching commitment to PSHE, or implemented a broad set of subject matter within non-evaluated PSHE curricula. Partly in recognition of the importance of the subject matter and partly in response to the acknowledged poor performance of schools in this area the UK government announced the intention to make PSHE education statutory in England in 2017, and compulsory from 2020 (UK Department of Education, 2019; House of Commons, 2019). The focus is to be on healthy relationships, health and well-being, safety and sexual health, but no curriculum guidance was given.

With a view to improving the haphazard and idiosyncratic approach adopted in teaching PSHE through defining an evidence based, integrated curriculum, a systematic review was undertaken by Coleman et al (2011). The aim was to provide evidence on the effectiveness of individual programmes relevant to PSHE education in secondary schools. Individuals involved in PSHE programme development, evaluation, implementation and education from Australia, the UK and the USA were also interviewed. Having established criteria for their evidence search and for module evaluation, implementation and delivery the researchers produced recommendations on a group of individual modules that could be combined to offer a comprehensive, cohesive package to deliver PSHE teaching.[6] This recommended package formed the basis for the *Healthy Minds* curriculum and programme of teaching, as based on an identified 14 individual teaching modules (see Coleman et al, 2011 for details). The individual modules selected to form the basis of the *Healthy Minds* curriculum are documented in Figure 1.[7] These modules have all been assessed on an individual basis, although generally within the USA rather than the UK, and concluded to have positive impacts across a range of domains on pupil's HRQoL. The individual modules were subsequently assessed with the objective of being combinable into a four-year PSHE curriculum to address issues of emotional health and wellbeing; diet/nutrition and healthy lifestyle; drugs, alcohol and tobacco education; safety education; and sex and relationships (Coleman et al, 2011).

2.1 Experimental Approach and Data:

In order to evaluate the defined *Healthy Minds* curriculum, it was rolled out as a clustered randomised trial. The sample size calculation to estimate the number of schools required for the trial started with the observation that the average English school has approximately 150 pupils per year. To allow for absentees and pupils leaving the school over the course of the trial we based our calculations on 100 pupils per year school cohort. With conventional statistical significance of 0.05 and power of

[6] In their meta-analysis of PSHE type teaching Durlak et al (2011) found that a positive outcome was a function of how well-designed and integrated the programme was. The latter point being especially important as short-term, non-integrated PSHE type teaching was found to be prone to fading effects (see also Bond and Hauf, 2004; Challen et al, 2011; Brunwasser et al 2009).

[7] See Appendix A Table A.1 for a figure which further illustrates the flow of lessons across each year of *Healthy Minds.*

0.80, and assuming a conservative intra-class correlation of 0.06[8] and equal numbers of treatment and control schools, (schools rather than school cohorts to allow a conservative estimate), a sample size of 25 schools was required to detect a treatment effect size of a change of 0.3 standard deviations in the HRQoL primary measure of self-assessed health, and at 30 schools we could detect an effect size of a change of 0.28 standard deviations.[9]

Randomisation to intervention or control group was based on the actual number of recruited school cohorts as described below. Minimisation of school differences was undertaken through identifying schools according to a number of national indicators that matched the intention of recruiting schools with poor attainment in above-average areas of deprivation.[10] Randomisation was undertaken by allocating a unique identifier to each school and use of a random number generator routine in Excel, with schools randomly allocated to 0 (control) or 1 (treatment), so that the randomisation process mimicked the flipping of a coin.[11]

For the school cohorts randomised to the *Healthy Minds* curriculum, the school teachers delivering the curriculum were given explicit training in the delivery of the curriculum. The training covered 7 days of teacher training for Year 7, (the entry year of pupils to secondary schools); 6 days of teacher training in Year 8, 2 days of teacher training in Year 9 and 4 days of teacher training in Year 10 (the year that pupils sit GCSE academic exams and choose to continue academic schooling or vocational schooling). Training covered all aspects of the 14 modules. This training was

[8] ICCs were reported to lie between 0.03 and 0.06 for comparable studies (Challen et al, 2011, UK Resilience programme evaluation: final report.).
[9] This treatment effect size is consistent with estimated standardised mean differences found in a number of studies of school interventions supported by mindfulness programmes, some similar in nature to a sub-set of the interventions proposed by the intervention under consideration, as well as cognitive behavioural interventions. Given that we might expect to have at least as great an impact through the *Healthy Minds* programme the treatment effect size was assumed to be approximately a change of 0.25 to 0.3 standard deviations. Hattie (2011, 2015, 2018) reports meta-analyses of proven effect sizes of various interventions in schools, documents these effect size findings and sets them within a wider context of school interventions.
[10] The criteria used were whether the percentage of pupils eligible for Free School Meals (FSM) was less than 13 per cent, between 13 and 25 per cent or greater than 25%; whether the percentage of pupils with 5 GCSEs with grades A*- C is below 59 per cent or not; and whether the school is single sex or mixed.
[11] School Recruitment encompassed a pragmatic element as school recruitment proved time consuming and complex Schools had to be recruited prior to first-year entry (September each year). Early contact was necessary, particularly given summer vacations. Unfortunately, some schools agreed early in the calendar year, merely to drop-out (for various reasons) as September approached (see Figure 2).

provided off-site by an independent charity, Bounce Forward[12], and covered both the material, and appropriate teaching methods for the *Healthy Minds* curriculum. The intervention, therefore, encompasses the curriculum content and teaching training and delivery.

The 4-year intervention trial began in 2013 (in Phase 1 recruited school cohorts). The *Healthy Minds* curriculum was delivered to children in the randomised intervention cohorts at age 11-12 years when they entered school year 7, their entrance year to secondary school. The curriculum was then taught as a 113-hour universal programme delivered over the first 4-years of secondary schools in the randomised sample of English classrooms, using one hour-a-week of timetabled lessons, replacing whatever non-standardized PSHE that had been historically timetabled for the same cohort.

In an effort to minimise dropout the initial control schools were recruited on a wait list control basis. For these wait list control schools, their school-year 7 entrants in 2013 acted as a control group, and their school-year 7 entrants in 2014 received the treatment.[13] Phase 2 (in 2014) recruited pure treatment and control school cohorts (i.e. to the end date of the study the control schools of Phase 2 did not received the treatment). Overall, the study recruited 13 participating schools in Phase 1 (2013), with 6 allocated to the (wait list) control arm and 7 to the treatment arm, and 21 participating schools in Phase 2 (2014), with 10 allocated to the control arm and 11 schools to the treatment arm. This gave a total of 34 schools, and, given the wait list control design, 39 school-cohorts. A participant flow diagram is provided in Figure 2.[14]

[12] The charity, Bounce Forward, also aided in the logistics of running (but not the analysis) of the trial within the schools.
[13] School-year 7 is the first (entrance) year to secondary (high) school in England.
[14] In Figure 2 the first panel highlights that 40 schools were approached and 37 agreed to participate. The second panel documents the original allocation of schools in each arm, and the number of pupils who are registered as being in these schools in total. The third panel details how many in each arm were present for the two follow ups. The final panel details how many questionnaires were handed in by pupils at each phase, and how many were valid (in this case they had a valid student number attached by the design team). The difference between the number of pupils in allocation and the questionnaires that were received back is owed to student absence, pupils choosing not to hand in their survey, as well as potentially missing responses.

Retention problems led to dropout over the course of the study. This is evident in the sample population differences reported in Figure 2 when going from allocation to follow up. Over time some schools were unable to maintain the teaching commitment. Some others were unable to provide support for the questionnaire administration which formed the basis of the data collection. Interim data collection, undertaken two years into the study, was therefore completed for 25 school cohorts only. A claw-back (re-engagement) mechanism was initiated for final year data collection (at year four of the study), where schools which had dropped out over the study period were contacted and asked if they were willing to participate in final data collection, and subsequently 35 school-cohorts were included in the final analysis.[15] We therefore focus on studying the impact of *Healthy Minds* identified from pooling the baseline and final data only using these 35 school-cohorts.

We also note that those classified as forming the treatment group, subsequently did not necessarily administer the curriculum in its entirety. Initial analysis therefore estimates the average treatment effect on an intention to treat basis incorporating attrition. As we have full details on compliance, we also estimate average treatment effects on the treated. See Table 1 for full details of the study timelines and Table 2 for full details of the number of schools who were present at each point of the study phase.

2.2 Data and measurement.

Data collection was carried out through questionnaires issued to individuals and conducted on school-sites at baseline (September 2013 or 2014 depending on school recruitment), 21 months (June 2015 or 2016) and 42 months (June 2017 with the final questionnaires delivered during 2018). Individual questionnaires were completed under standard exam conditions within the individual schools, with each session lasting 1-hour. Participants were informed at the start of the session that the survey data would be collated anonymously, and that parents, teachers or other pupils would

[15] There was an additional special-needs school which participated in the study, but as it did not meet the inclusion criteria it was excluded from the final analysis.

not have any access to the data. Anonymous coding ensured that the analysts also do not have access to individual pupil's identities.[16]

There is no single, comprehensive measure that can capture the full range of soft skills (Conti and Heckman, 2012; Decancq and Neuman, 2014; Kahneman and Krueger, 2006). Our outcome measure was taken from a well validated questionnaire, the CHQ-CF87 (Schmidt, Garratt and Fitzpatrick, 2001). The CHQ-CF87 has been found to be reliable and sensitive to measured changes across a range of domains for 10-18 year-olds (Schmidt, Garratt, and Fitzpatrick, 2001). The questionnaire is therefore suitable for and has been validated within a school context and takes an estimated 20 minutes to complete (see Schmidt, Garratt, Fitzpatrick, 2001). It is based on 87 items that measure physical and psychosocial health, divided across 14 multi-item scales on physical functioning, socio-emotional role, social-behavioural role, social-physical role, pain, general behaviour, mental health, self-esteem, general health perceptions and family activities.[17] A list of the sub-scales and their number of associated items is documented in Table 3.[18]

While the presence of multiple outcomes (scales) does lead to the potential problem of selective reporting and under-reporting of heterogenous effects, the CHQ-CF87 incorporates a single scale relating to self-assessed health and this defined the primary outcome of the trial and formed the basis of the ex-ante calculation for the study sample size as reported above. This self-reported health measure, corresponds to the following text "*In general, how would you say your health is*?" and recorded across 5 levels ranging from bad to excellent. The remaining CHQ scales are treated as secondary outcomes.

The CHQ-CF87 coding manual details the approach to deal with missing values. Specifically, most of the instruments are multiple items (see Table 2), so in cases

[16] Questionnaires were administered, collected, collated, and the data coded by an independent firm (HcareSolutions). Anonymised data were forwarded to the analysts in Excel and analysed in Stata15.
[17] Unfortunately, the one question that captures the scale relating to past health was omitted from the baseline questionnaire (it was dropped by the company commissioned to print the questionnaire in error). So, our study relies on thirteen CHQ-CF87 scales on which to assess impact.
[18] For full details of the sub scales and their associated items please see https://www.healthactchq.com/survey/chq. Our license to use the CHQ-CF87 does not allow us to reproduce these items here.

where one answer is missing, the total is imputed from the answers that are returned with positive values, assuming a pre-specified number of questions in the instrument have been answered.[19] To give balanced samples across our reported regressions restrict samples only to where the primary outcome question is answered. However, an Appendix (B) documents all of the estimates presented in main text tables absent of this restriction. None of the point estimates are discerningly different.

3. Methodology

The primary analysis is an intention to treat analysis. This is presumably the treatment effect policy makers care about the most, as if the curriculum is adopted there will be heterogeneity in how it is rolled out at the school level. Of course, if the value placed on this curriculum increases in the UK over time (i.e. if this curriculum is adopted over time within the compulsory PSHE teaching introduced by UK government) we might expect that compliance issues to somewhat resolve, with a consequent rise in the treatment effect. In other words, given statutory support for the *Healthy Minds* curriculum we expect that our estimated effects are conservative.

The primary empirical analysis specified ex ante[20] is aimed at recovering the average treatment effect and based on the following basic difference-in-difference specification:

$$y_{ist} = \beta_s + \beta_1 treatment_s + \beta_2 year_t + \beta_3 treatment * year_{ts} + \epsilon_{ist} \quad (1)$$

where: y_{ist} is the CHQ-CF87 primary outcome variable (self-assessed health), β_s is a set of school fixed effects, $treatment_s$ is equal to 1 if a school was randomised to treatment, regardless of whether they adhered to the treatment and zero otherwise, $year_t$ is a set of yearly dummy variables based on the year the data was collected. Equation (1) is subsequently estimated with the baseline and the final year data, including the clawed-back data. The coefficient β_3 captures the average effect of being randomly assigned to the treatment with an intention to deliver the outlined

[19] For a very small number of single items (there are two including the primary outcome) if required imputed values were based on the average given in the school at that data collection point.

[20] This was submitted as part of a statistical analysis plan to the Education Endowment Foundation who partially funded this work.

PSHE curriculum. The standard errors, estimated using wild bootstrapping are clustered at the level of the pupil within a school[21].

Similar regressions are run for the other CHQ-CF87 secondary outcome scales. This raises the issue of multiple hypothesis testing and therefore, we document the probability of significance after applying a correction for multiple testing. The approach adopted is that proposed by Benjamini and Hochberg (1995) based on testing for the false discovery rate, as calculated on the proportion of falsely rejected hypotheses in a joint test of a set of hypotheses, to adjust for the effect of multiple testing.[22]

We also collected detailed compliance information throughout the project. Compliance is defined through identifying the schools stopping point in the curriculum delivery in accordance with Figure 1. We calculate the proportion of the *Healthy Minds* curriculum that a particular school cohort achieved, through each year counting as 0.25 completion of the intervention, given the 4-year length of the trial. This completion rate ranges from 0 to 1, and using this information we estimate:

$$y_{ist}=\beta_s+\beta_1 treatment_s+\beta_2 year_t+\beta_3 compliance*year_{ts}+\varepsilon_{ist} \quad (2)$$

In Equation (2) $compliance$ denotes the proportion of the *Healthy Minds* programme completed by an individual school (*s)* in the final year (in essence it is $compliance$ interacted with a dummy that denotes the post treatment period) , all other variables defined as for equation (1). We expect that there are systematic differences between the schools who comply and those that do not. Therefore, we retrieve β_3 as an average treatment effect utilising assignment to treatment as an instrumental variable. Unsurprisingly, this instrument is strong (F of first stage is 125), and given that assignment to treatment is randomised by the research team we are also certain of the instrument's validity.

[21] In Appendix F Table F.1 we document our main analysis with standard errors clustered at the school level.
[22] The Benjamini and Hochberg (1995) article outlines that this approach is more robust than that based on the familywise error rate (FWER).

As indicated from Table 2, a number of pupils dropped out of the study, rendering their outcomes in the subsequent follow up phases unobserved. A second robustness allows for this by first estimating the probability of dropout among continuing individuals to construct inverse-probability weights (IPWs) to adjust for adherence. Given that we do not observe student demographics in the initial data collection, these weights are constructed by regressing a discrete variable, indicating whether a student remained in the study or not, on their initial CHQ-CF87 scales and a set of school fixed effects. These weights are then applied when re-estimating Equation (1).

A unique identifier was also assigned to pupils who answered questionnaires as part of *Healthy Minds* data collection, allowing identification of pupil fixed effects. Given the systematic differences identified at the pupil level in Table 4, we consider a model that includes these pupil fixed effects as a robustness test. That is, we estimate:

$$y_{ist} = \beta_i + \beta_1 treatment_s + \beta_2 year_t + \beta_3 treatment * year_{ts} + \epsilon_{ist} \quad (3)$$

where β_i is an individual fixed effect, and again all other variables defined as in Equation (1). Given that treatment assignment is random, it is likely that *treatment*year* is orthogonal to the random noise term, and also the individual fixed effects. Given that we are interested in estimating β_3, this implies that a random effects model is appropriate. However, in applying this equation across the range of outcomes, we test this explicitly using the Hausman test and estimate through fixed effects in the few cases where random effects are not supported.

4. **Results**

Our main estimation sample is based on 23 treatment schools (including wait list control schools clawed-back for final analysis) and 19 controls. These schools provided 39 school cohorts and a total of 3,789 school pupils with valid primary outcome data at the point of randomisation. At the end of the study, allowing for clawback after drop out, 35 school cohorts formed the basis of the final analysis and 3,537 school pupils (2,236 in the treatment arm and 1,301 in the control arm) for which comparative data were available over the complete period of the study. This gives a total of 7326 valid questionnaires for analysis, implying that this number of respondents answered the primary outcome single item question. Our balanced

analysis (with imputation providing that a certain number of questions are answered in multi item questions) is based on 7252 pupils. This is made up of 3789 pupils at the point of randomisation and 3463 at t=4. The difference is caused by pupils not answering any or too few of the multi-item questionnaires.

Table 4 documents the unadjusted differences in the mean scale levels for the items in the CHQ-CF87 questionnaire between the treatment and control schools at baseline for the total of 3,789 pupils. Scales have been standardised to have a mean of zero and standard deviation of one. A negative sign denotes that the average treated child had a worse outcome at baseline, conversely a positive sign denotes that they were better off. Table 4 also documents standard errors for these differences in brackets. As can be seen, for five of the thirteen outcomes, there are no significant differences between the average treated and control child. This includes the primary outcome (global health) scale. However, significant differences are observed in the eight remaining outcomes, with the children assigned to the intervention, the *Healthy Minds* curriculum, being notably worse at initiation in the physical functioning, behaviour, global behaviour, mental health, self-esteem and general health scales. The children allocated to the intervention are very slightly better off in the family activities and family cohesion scales. Our modelling strategy pursued through Equation (1) assumes that without the *Healthy Minds* curriculum these differences would remain fixed. Other analyses (notably the pupil effects analysis described in Equation 3) consider alternative approaches, which relaxes this assumption.

Table 5 documents the main results from our standard difference in difference models (see Equation 1). All of the reported estimates are given with the statistical significance at conventional levels unadjusted for multiple comparisons. Given that the primary outcomes are part of 13 scales analysed from the CHQ-CF87 questionnaire adjustment was also made to allow for multiple comparisons. These are reported in the rows denoted by MCC (Multiple Comparison Confidence intervals), and reflect the Benjamini and Hochberg (1995) corrections.

As can be seen for the primary outcome global health, reported in Table 5, the effect is positive and significant in the baseline result. Pupils exposed to *Healthy Minds* have global health attainment that is 0.235 standard deviations higher than children in

the control group. This effect is significant at the 5% level of significance based on the Benjamini and Hochberg (1995) multiple comparison correction (B&H in Table 5 and from herein). This aligns well with findings by Durlak et al (2011) that typical programmes raise outcomes by around 11 percentile points: our estimates imply that outcomes are raised by 10 percentile points. The estimated coefficient is also consistent of a general finding that more than 60% of individuals in the treatment group return a self-assessed general health improvement, which is above that of the control arm individuals as a result of the intervention. Given that *Healthy Minds* is given to adolescents, and most of the programs in the meta-analysis are given to younger children, we view this as evidence that adolescent's HRQoL is malleable to similar levels throughout childhood.

Additionally, from Table 5 there is evidence of significant, positive and substantive gains to a number of the secondary outcomes. These are: physical difficulties, pain and discomfort, family activities and family cohesion scales to the order of 0.288, 0.239, 0.174 and 0.242 respectively. We note there is a negative effect for self-esteem, albeit it is not significant when we draw on traditional t-testing or the B&H correction. Still, the estimate is relatively substantive[23]. Given that this is attributable to the intervention it could be indicative of greater self-reflection amongst adolescents as a result of receiving the intervention. We also note that all other emotional wellbeing outcomes are positive but centred close to zero (emotional difficulties and mental health).

The top panel of Table 6, presents the our 2SLS estimates (Equation 2), which recover an average treatment effect on the treated by using assignment to *Healthy Minds* as an instrument for a variable which measures the level of compliance with the program. As expected, the estimates in Table 6 are larger for all outcomes (except global behaviour) as compared to Table 5 (i.e. utilising the intention to treat approach biases the effects downwards given that not all schools complied perfectly so there are never takers in the treatment pool). The primary outcome, global health attainment now improves by 0.296 standard deviations in the intervention group as compared to the control. We also have a good concordance with the conclusion that we have

[23] We note that this effect is significant when we consider some robustness analysis (see below) and robustness in clustering (see Appendix F).

drawn so far (physical difficulties, pain and discomfort and family activities are all significantly augmented by *Healthy Minds.* The effect for family activities remains positive and substantive but is now not significant). We note in Table 6 self-esteem remains negative and is now significant.

The bottom panel of Table 6 presents the estimates emanating from applying IPW weights adjusting for adherence to Equation 1. These estimates as compared to Table 5 change little, with the majority of the effects attenuated slightly. The IPW estimates imply that the primary outcome, global health attainment, improves by 0.212 standard deviations in the intervention group as compared to the control.

Table 7, presents our estimates from Equation 3, which add a pupil effect to the baseline model. The Table also details the estimator that is used (random effects versus fixed effects), and also the p-value emanating from the Hausman test which allowed us make this decision.[24] We note that the effects are again comparable with what has gone before, i.e. most estimates are very similar to those shown in Table 5.

4.1 Summarising the Effects

The Tables 5 through 7 document estimates from four different models for thirteen different outcomes. In order to condense the takeaway message from this study we reduce the dimensionality of twelve of these outcomes (we exclude the primary outcome from the dimensionality reduction), into three aggregate variables that are intuitively labelled physical health, emotional wellbeing and behaviour. To achieve this, we choose the items in the CHQ-CF87 secondary outcomes, which directly relate to these three distinct domains. Specifically, physical health comprises of the individual domains of physical health, family activities, general health and physical functioning. Emotional health comprises the individual domains of self-esteem, mental health, and emotional difficulties. Finally, behaviour comprises of behavioural

[24] Here the null is that the random effects is preferred.

difficulties, global behaviour, family activities and behaviour.[25] These variables are combined into these three distinct HRQoL aggregates using exploratory factor analysis (for full details see Appendix C). Figures 3a through 3d then summarize our findings across these three latent factors, alongside the primary global health outcome, as a summary of the overall effects of *Healthy Minds*.

Overall, Figures 3a through 3d present clear evidence that *Healthy Minds* augments global health and physical health. The effects for behaviour are also substantive and consistently significant at the 10% level of significance. The overall impact on emotional wellbeing is zero across all four models.

4.2 Differences by Gender[26]

We note that a related evolving literature highlights that non-cognitive skills are the basis of improved outcomes witnessed in girls, such as improved school and college attainment. This raises the question of whether investment in such skills might have bigger gains to boys. For example, Jacobs (2002) and Goldin, Katz and Kuziemko (2006), document that boys tend to have great disciplinary problems in school and in later life, and invest less time in homework. Indeed, Goldin, Latz and Kuziemko (2006) find that controlling for non-cognitive traits eliminates female college attainment advantage, raising the possibility of a differential gender impact in non-cognitive skill investment. The importance of exploring the potential for heterogenous treatment effects by gender *in Healthy Minds* is bellied in the already realised gap in attainment between boys and girls in the UK at age 12, when *Healthy Minds* begins (see Table 8).[27] Additionally, other programs aimed at augmenting soft skills have found differential effects by gender. See for example the Perry Preschool

[25] We note that family activities enters both the health and behavior latent factors because the question specifically relates to 'how often has your health and behaviour' negatively affected family activities in six distinct ways.

[26] We note that we only ask for gender when t=4 and it is asked at the end of the questionnaire causing significant non response. We therefore replicate all the results from the main tables in Appendix D withthis smaller sample.

[27] The overall message from Table 8 is that at the time of *Healthy Minds* initiation males have lower starting points on behaviour, but higher starting points on emotional wellbeing. We note that these differences are also realised in the general adolescent population of the UK. With respect to health the observed initial gender differences vary across the CHQ sub scale chosen, with males having a worse starting point on the primary outcome (global health) as well as physical functioning and physical differences. However, they have a better starting points on pain and discomfort.

program (Heckman, Pinto, and Savelyev (2013)) and The Abecedarian Program (Campbell, Conti, Heckman, Moon, and Pinto (2013)).

We therefore explore the potential for heterogenous treatment effects by gender in *Healthy Minds*. To do so we add a male dummy variable and an interaction between this dummy and the treatment effect to the pupil effects analysis described in equation (3). This allows examination of whether *Healthy Minds* changes the HRQoL outcomes of males and females differently, while continuing to control for unobserved individual pupil effects.

Table 9 documents the estimates from the additional analyses which explores heterogenous treatment effects by gender. It is notable, that *Healthy Minds* augments many more outcomes of male pupils significantly and substantively[28] (global health, physical functioning, emotional difficulties, behavioural difficulties, self-esteem, pain and discomfort, behaviour, mental health, general health, family cohesion and family activities) as compared to females (global health, pain and discomfort and family cohesion). Notably, there are also significant and negative effects for females in the following domains: physical functioning, self-esteem, mental health and general health. Overall, we conclude, that teaching HRQoL in mixed sex schools is not straightforward, as we interpret the evidence from Table 9 as highlighting that boys and girls learn these skills differently and this results in differential responses to the *Healthy Minds* curriculum.

4.3 Labour Market Outcomes:

The evaluation data ends with the final data collection in Year 10, and given the majority of children remain in school within our analysis period, we cannot follow the pupils into the labour force. However, we did collect data on the child's occupational aspirations. Specifically, in the final year of the study the participants are asked to identify the job they aspire to when they are 30 years old. We can relate this occupation to a 3-digit occupation code, which in turn allows us to relate it to variables that proxy specific attributes of their chosen job. Specifically, we are interested in exploring the impact of *Healthy Minds* on occupational aspirations in

[28] As before we read significance from the B&H probabilities.

terms of the following characteristics (full details on the construction of these variables can be found in Appendix E):

1. *Share of males*: we are specifically interested in gender differences and whether *Healthy Minds* caused decreases in the number of boys choosing roles that are traditionally male and/or increases for girls and proxy this through the proportion of male employment at occupational level.
2. *Log Average hourly income*: we are specifically interested in examining whether the average treatment effect of *Healthy Minds* causes its pupils to become less extrinsically motivated through consideration of their aspirational income as related to their aspirational occupation.
3. *Average hours*: we are specifically interested in examining whether the average treatment effect of *Healthy Minds* causes its pupils to become more aware of work life balance, and thus choose occupations with lower average hours.
4. *Wage-hours elasticity* (as defined by Goldin, 2014): Goldin interprets this occupation specific elasticity as capturing the wage penalty arising from working shorter hours. An average treatment effect that is negative and significant implies that *Healthy Minds* is causing its pupils to shy away from occupations that are hours intensive and the best paid. Conversely, an average treatment effect that is positive and significant implies that *Healthy Minds* is causing its pupils to move towards these occupations.
5. *Competitiveness*: we are specifically interested in examining whether the average treatment effect of *Healthy Minds* causes its pupils to move towards jobs that are more collaborative, and thereby less competitive.
6. *Job content*: Following Lordan and Pischke (2016) we create three latent factors that capture the level of job interaction with '*people'* (i.e. dealing with people on a day to day basis), '*data'* (i.e. dealing with data and problem solving on a daily basis) and '*things'* (i.e. creating and fixing objects on a daily basis). White collar jobs that are relatively high on *people* include social work, teaching and nursing. Examples of jobs that are relatively high on *data* include financial managers, mathematicians and statisticians. Finally, for *things*, engineers and architects are relevant examples. We are specifically interested if *Healthy Minds* caused a change in the job content being aspired to by its pupils towards occupations that are more *people* orientated.

Given that we only have data for the final collection on occupational aspirations we estimate the impact of *Healthy Minds* on these outcomes as follows:

$$y_{is} = \beta_s + \beta_1 treatment_s + \epsilon_{is} \quad (4)$$

All definitions are consistent with Equation 1, only in this case *y* is a proxy that represents an aspect of a child's occupation aspirations. Given that the identification strategy requires the treatment and control group to be strictly comparable to infer causality (rather than allowing for fixed differences over time as for our other estimates), and the established baseline differences documented in Table 10, we present estimates for Equation 4 tentatively. We complement these estimates with a model that adds a male dummy, and an interaction between this dummy and $treatment$, allowing us to identify whether the estimates obtained differ by gender.

From Table 10 panel 1, there is highly suggestive evidence that *Healthy Minds* affects aspirations. Specifically, the estimates suggest that pupils exposed to the treatment choose jobs that have 10% lower share of males. The effects on weekly income are also negative and significant, but not substantive; earnings are £2.70 less per week. Notably, those exposed to treatment are less likely to seek out competitive work (0.347 standard deviation decrease) and far more likely to choose jobs that involve *people* interactions (0.283 standard deviation increase).

The bottom panel of Table 2, suggests that the estimates for income, *people* and competitiveness are sample average effects: i.e. not attributable to any specific gender, in almost all cases. This is evidenced by the coefficients on the male interaction terms being not significant and centered around zero. The exception is for the variable associated with occupations (higher) shares of males. Interestingly, this estimate implies that *Healthy Minds* significantly decreases the likelihood that male pupils will choose traditional male work. Additionally, the overall effect is zero, implying male pupils alone are affected. The *people* coefficient in the bottom panel of Table 2 remains substantive, but noisy standard errors render it not significant. The estimates still imply significant shifts away from competitive work because of *Healthy Minds.* In addition, the estimates suggest that pupils treated with the program

more often seek out jobs that are high on '*things*', but less often seek out jobs that are high on '*data*.'

4.4 Overall Cost Effectiveness

Table 9, details the costs of delivering *Healthy Minds.* The teacher training costs associated with delivery of the *Healthy Minds* curriculum total £3,600 per school over the 4-years of the programme. Including replacement teacher costs these almost double, to £6,640 per school. Including reasonable estimates for materials the total cost of delivery over the 4-year programme run to £7,250 per school or £1,812.50 per school per year. Or assuming a single teacher teaches the curriculum to 3 classes per year of 30 pupils, this is approximately £20.10 per student taught per year. According to Belfield, Crawford and Sibieta, (2018) secondary school spending per pupil is estimated to be £6,000 in 2018/19, so the Healthy Minds programme would add under 0.4% per pupil expenditure per annum.

A further way to put this into context is to consider a conversion into the widely adopted Quality Adjusted Life Year (QALY) gained outcome metric used in the UK health care sector. Taking our reported result of an increase of 0.235 standard deviations in pupil's self-reported general health as our preferred measure of primary outcome, we assume this to represent a proxy measure of QALY gain across a [0,1] scale.[29] We then estimate a pessimistic QALY gain attributed to the *Healthy Minds* curriculum as 0.055 QALYs gained (calculated as 0.235 * 0.235), while an optimistic estimate attributed to the treatment could be taken as 0.235 QALYs gained over the 4-year programme. If this was achieved at £100 per pupil (in fact, if the per pupil per year cost were £20.10 as calculated above it would be achieved at £80.40 per pupil taught the 4-year curriculum) then the cost per QALY achieved through the taught *Healthy Minds* curriculum would range between a pessimistic estimate of approximately £1,811 per QALY and an optimistic estimate of £426 per QALY gained. This range is well below the lower limit of £20,000 to £30,000 per QALY gained operated by the UK's National Institute for Health and Care Excellence (NICE) when considering the introduction of new health care treatments into the

[29] QALYs are normally transformed to lie between a scale of 0, worst imaginable health state (normally calibrated to death) and 1, best imaginable health state (normally associated with "perfect" health).

English National Health Service (NICE, 2013). It is clearly a cost-effective public health intervention as per NICE guidelines on these grounds.

5. Conclusions

Our work evaluates, through a randomised trial, an evidence based HRQoL curriculum called *Healthy Minds* that ran in 34 schools over a four-year period. We focus on the impact *Healthy Minds* has on health, behaviour and emotional wellbeing as measured by the validated CHQ-87 instrument. Given recruitment and compliance issues the trial was analysed on an intent-to-treat basis. On this basis we find that *Healthy Minds* increases the primary outcome of this study – global health – by 0.235 standard deviations. This is consistent with more than 60% of individuals in the treatment arm returning a self-assessed general health improvement above that of the control arm as a consequence of the curriculum.

In addition, we find robust evidence that *Healthy Minds* positively augments many other physical health domains. Finally, a summary of our results presents some evidence (significant at the 10% level of significance) that *Healthy Minds* positively affects behaviour but has no impact on emotional wellbeing. We note that the impact on self-esteem is negative and significant in some specifications. The cause of this is not known and warrants further consideration. However, we have speculated that through raising self-awareness generally at a time of adolescence, this might account for this specific finding, and therefore although out of line with the generality of positive findings associated with the *Healthy* Minds curriculum, may not necessarily be a wholly detrimental outcome. These, conclusions are robust to a number of alternative specifications, which allow for known analytical difficulties associated with randomised trials. These include randomisation issues (dealt with through an intent-to-treat analysis); multiple hypothesis testing (Benjamini and Hochberg adjustment); compliance (IV analysis) and adherence (Inverse Probability Weighting).

An analysis that allows for heterogenous treatment effects by gender highlights that boys have significant gains in many more outcomes as compared to girls. Differential effects by gender are commonly found in programs that aim to teach soft skills, but it does raise a question on how these programs should be taught to ensure that

everyone is a 'winner' in mixed gender schools. Again, this is an avenue for future research as our study does not indicate why these differences emerge.

We also explore the effects of *Healthy Minds* on career aspirations and find that treatment with *Healthy Minds* predicts the likelihood of choosing competitive work negatively, but positively predicts the likelihood of choosing work that is high on '*people*' content. This suggests that *Healthy Minds* is causing pupils to prefer work that is more social and collaborative, albeit the identification strategy is not watertight enough to have confidence that the estimate is causal. It is also an aspirational desire and may, of course, not be realised.

We note that our recruitment strategy principally targetted school populations which were disadvantaged and it remains an open question whether our principle measure of non-cognitive skills (health, behaviour and certain internal feelings) are more heterogeniously disributed within such schools as compared to more advantaged school populations. Given that resource constrains are arguably more binding in the schools in our study, our selection strategy may also have contributed to the witnessed compliance difficulties. It is therefore possible that if the *Healthy Minds* curriculum was introduced on a compulsory basis stronger impacts would be achieved.

Our work provides evidence of the possibility that HRQoL, and career aspirations, can be changed through a school based initative after the age of 11. Our work is important as it is the first study of an evaluation on this adolescent age group which considers an intervention specifically designed to change non-cognitive skills. It is also the first study of its kind in a UK setting. This reflects the thin literature on the impact that schooling has on adolescent outcomes generally. Durlak et al (2010), for example, in a meta-analysis of after-school programmes (APS) aimed at improving personal and social skills, explicitly comments on the small number of studies reviewing impacts on adolescents. There is simply, little knowledge over the impact of interventions aimed at adolescents; the "missing middle" as Almond, Currie and Duque (2018) call it. Even where adolescents are studied and elements of non-cognitive skills are analysed, attention is normally directed to schooling performance and employment outcomes rather than on their health or behaviour (e.g. see Kautz, T. Heckman, J., Diris, R., et al., 2014).

Currently there is a debate in the UK around specific governmental policy on PSHE that encourages investments in the knowledge, understanding and non-cognitive skills adolescents need to manage their lives (House of Commons Education Committee, 2015; PSHE Association, 2017). The debate in England has culminated with the UK government introducing guidance in 2019 that obliges secondary schools to timetable a PSHE curriculum in 2020. The importance of these skills is echoed in the growth in their need on the labour market in the UK because of an increasing tendency for jobs to be automated that do not need these skills. Independently, preparing teenagers for a healthy and happy life is important in and off itself. Our work highlights that this can be done with some success at the school level. Future research can consider alternative curriculums so we can learn more about how HRQoL should be taught, and what should be avoided, to maximise its outcome.

References

Acemoglu, D. (2002), 'Technical change, inequality, and the labor market', Journal of Economic Literature, 40, 7–72

Almond, D. Currie, J., and Duque, V., (2018), "Childhood circumstances and adult outcomes: Act II" Journal of Economic Literature, 56, 1360-1446

Autor, D. H., Dorn, D., (2013), "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." American Economic Review, Vol. 103, No. 5, August, pp. 1553-97.

Autor, D., Dorn, D., and Hanson.G., (2015), "Untangling Trade and Technology: Evidence from Local Labor Markets." Economic Journal, Vol. 125, No. 584, May, pp 621-46.

Autor, D. H., Levy, F. & Murnane, R. J. (2003), 'The skill content of recent technological change: An empirical exploration', The Quarterly Journal of Economics, 118, 1279–1333

Baker, M., Gruber, J. and Milligan, K., (2015), "Non-cognitive Deficits and Young Adult Outcomes: The Long-Run Impacts of a Universal Child Care Program." National Bureau of Economic Research Working Paper 21571.

Belfield, C., Crawford, C. and Sibieta, L., (2018), Long-run Comparisons of Spending per Pupil Across Different Stages of Education, Institute of Fiscal Studies (IFS), London

Bekman, E., Bound, J. & Machin, S. (1998), 'Implications of skill-biased technological change: international evidence', Quarterly Journal of Economics, 113, 1245–1279.

Benjamini, Y. and Hochberg, Y., (1995), "Controlling the False Discovery Rate: A practical and powerful approach to multiple testing", Journal of the Royal Statistical Society, Series B (Methodological), 57, 289-300

Bhargava, S., Sachdev, C.H., Fall, C.H., et al, (2004), "Relations of serial changes in childhood body-mass index to impaired glucose intolerance in young adulthood", New England Journal of Medicine, 350, 865-875

Black, S., Devereux, P., and Salvanes, K., (2007), "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes", The Quarterly Journal of Economics, 122(1): 409-439.

Bond, L. A., & Hauf, A. M. C., (2004), 'Taking stock and putting stock in primary prevention: Characteristics of effective programs', J. of Primary Prevention, 24:199–221.

Bonell, C., Fletcher, A. & McCambridge, J., (2007), “Improving school ethos may reduce substance misuse and teenage pregnancy”, BMJ, 334:614–16.

Brunwasser SM, Gillham JE, & Kim ES., (2009), “A meta-analytic review of the Penn Resiliency Program's effect on depressive symptoms”, J Consult Clin Psychol. 77(6):1042-54.

Campbell, F.A., Conti, G., Heckman, J.J., Moon, S.H., Pinto, R., Pungello, E.P. and Pan, Y., (2014), ‘Early childhood investments substantially boost adult health’, Science, vol. 343, 478–85.

Case, A. and Paxton, C, (2011), “The Long Reach of Childhood Health and Circumstance: Evidence from the Whitehall II Study”, Economic Journal, 121, F183-F204

Case, A. and Paxton, C, (2010), “The Long Reach of Childhood Health and Circumstance: Evidence from the Whitehall II Study”, NBER Working Paper 15640, Cambridge, NBER

Case, A., A. Fertig, and C. Paxson, (2005), “The lasting impact of childhood health and circumstance”, Journal of Health Economics 24, 365-389.

Challen, A., Machin, S , Noden, P., & West, A. (2011), UK Resilience Programme Evaluation: Final Report. Department for Education, Research Report DFE-RR097 .

Coleman, J., Hale, D. & Layard, R., (2011), A Model for the Delivery of Evidence-Based PSHE (Personal Wellbeing) in Secondary Schools. CEP DP 1071.

Collishaw, S., Maughan, B., Goodman, R. & Pickles, A., (2004), “Time trends in adolescent mental health”, J. of Child Psychology and Psychiatry, 45(8):1350-62.

Conti, G., Heckman, J.J. and Urzua, S., (2010), The Education-Health Gradient, American Economic Review (Papers and Proceedings), 100, 234-238

Conti, G. and Heckman, J.J., (2014), ‘The economics of child well-being’, in (A. Ben-Arieh, F. Casas, I. Frønes and J.E. Korbin, eds.), Handbook of Child Well-Being, pp. 363–401, Dordrecht: Springer.

Conti, G. and Heckman, J., (2016), “The effects of two influential early childhood interventions on health and behaviour”, Economic Journal, 126, F28-F65

Cunha, Flavio, and James J. Heckman., (2008), “Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation.” Journal of Human Resources 43 (4): 738–82.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010), “Estimating the technology of cognitive and noncognitive skill formation”, Econometrica 78 (3), 883{931.

Currie J., (2009), “Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood and Human Capital Development”, Journal of Economic Literature, 47(1): 87-122.

Currie J., (2011), “Inequality at Birth: Some Causes and Consequences”, American Economic Review, 101(3): 1-22.

Decancq, K., and Neumann, D., (2014), Does the choice of well-being measure matter empirically? An illustration with German data, IZA Discussion Paper 8589

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B., (2011), “The impact of enhancing pupils' social and emotional learning: A meta-analysis of school-based universal interventions”, Child Development, 82(1):405-32.

Durlak, J. A., Weissberg, R. P., & Pachan, M., (2010), “A meta-analysis of after-school programs that seek to promote personal and social skills in children and adolescents”, American J. of Community Psychology, 45:294–309.

Formby, E., Coldwell, M., Stiell, B., Demack, S., Stevens, A., Shipton, L., et al., (2011), Personal, Social, Health and Economic (PSHE) Education: A mapping study of the prevalent models of delivery and their effectiveness. Department for Education.

Garner, A.S., Shonkoff, J.P., Siegel, B.S., Dobbins, M.I., Earls, M.F., Garner, A.S., McGuinn, L., Pascoe, J., Wood, D.L. and Committee on Psychosocial Aspects of Child and Family Health, Committee on Early Childhood, Adoption, and Dependent Care, and Section on Developmental and Behavioral Pediatrics, (2012), ‘Early childhood adversity, toxic stress, and the role of the pediatrician: translating developmental science into lifelong health’, Pediatrics, vol. 129(1), pp. e224–31.

Goldin, C., Katz, L. and Kuziemko, I., (2006), “The homecoming of American college women: The revesal of the college gender gap”, Journal of Economic Perspectives, 20, 133-156

Goodman, A., R. Joyce, & Smith J.P., (2011), “The long shadow cast by childhood physical and mental problems on adult life”, PNAS, 108(15):6032-37.

Hattie, J..C., (2011), ‘Which strategies best enhance teaching and learning in higher education?’, in Empirical Research in Teaching and Learning, Mashek, D. and Hammer, E., (eds.), (Blackwell, London)

Havnes, T., and Mogstad, M., (2015), “Is Universal Child Care Leveling the Playing Field?”, Journal of Public Economics 127: 100–114.

HealthActCHQ, (2015), The CHQ Scoring and Interpretation Manual, Boston:MA

Heckman, J. J., Pinto, R., and Savelyev, P., (2013), “Understanding the mechanisms through which an influential Early Childhood Program boosted adult outcomes”, American Economic Review 103 (6): 2052–86.

Heckman, J. J. and Kautz, T., (2014), 'Fostering and measuring skills: Interventions that improve character and cognition', In J. J. Heckman, J. E. Humphries, and T. Kautz (Eds.), The Myth of Achievement Tests: The GED and the Role of Character in American Life, pp. 341–430. Chicago, IL: University of Chicago Press.

Heckman, J. J., R. Pinto, and Savelyev, P.A., (2013), 'Understanding the mechanisms through which an influential early childhood program boosted adult outcomes", American Economic Review 103 (6), 1–35.

Heijmans, B.T., Tobi, E.W., Stein, A.D., Putter, H., Blauw, G.J., Susser, E.S., Slagboom, P.E. and Lumey, L.H., (2008)., "Persistent epigenetic differences associated with prenatal exposure to famine in humans", Proceedings of the National Academy of Sciences, vol. 105(44), pp. 17046–9.

Hoeschler, P., Balestra, S., and Backes-Gellner, U., (2018), "The development of non-cognitive skills in adolescence", Economics Letters, 163, 40-45

House of Commons Briefing Paper, (2019), Relationships and Sex Education in Schools (England), Briefing Paper 06103, UK Parliamentary Papers.

House of Commons Education Committee, (2015), Life lessons: PSHE and SRE in schools. Fifth Report of Session 2014-15. HC 145. London: HMSO

Humphries, J. and Kosse, F., 2017, On the interpretation of non-cognitive skills – What is being measured and why it matters, Journal of Economic Behaviour and Orgnaisation, 136, 174-185

Humphrey, N., Lendrum, A., & Wigelsworth, M., (2010), Social and emotional aspects of learning (SEAL) programme in secondary schools: National evaluation. Department for Education.

Jacob, B., (2002), 'Where the boys aren't: Non-cognitive skills, returns to school and the gender gap in higer education", Economics of Education Review, 21, 589-98

Jayachandrand, S. and Pande, R., (2017), "Why are Indian children so short? The role of birth order and son preference", American Economic Journal, 107, 2600-2629

Kahneman, D., and Krueger, A. B., (2006), "Developments in the measurements of subjective well-being", Journal of Economic Perspectives, 20, 3-24.

Kautz, T. Heckman, J., Diris, R., et al., (2014), Fostering and Measuring Skill: Improving cognitive and non-cognitive skills to promote lifetime success, NBER Working Paper, 20749, Cambridge, NBER

Kniesner, T. J., and ter Weel B., (2008), "Special Issue on Noncognitive Skills" Journal of Human Resources, 43, 729-1059, [884]

Klein JB, Jacobs RH, Reinecke MA., (2007), "Cognitive-behavioral therapy for adolescent depression: a meta-analytic investigation of changes in effect-size estimates", J Am Acad Child Adolesc Psychiatry, 46: 1403–13

Lekfuangfu W, and Lordan G (2018). Cross Cohort Evidence on Gendered Sorting Patterns in the UK: The Importance of Societal Movements versus Childhood Variables. IZA DP No. 11872

Lleras, C., (2008), "Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings", Social Science Research 37 (3), 888–902.

Lordan Grace and Pischke Jörn-Steffen (2016). Does Rosie Like Riveting? Male and Female Occupational Choices. NBER Working Paper No. 22495. Issued in August 2016. NBER Program(s): LS.

Lordan, Grace and Neumark, David (2018). People versus Machines: The impact of minimum wages in automatable jobs. *Labour Economics.* 52 (June), pp 40-53.

NICE, (2013), How NICE measures value for money in relation to public health interventions, Local government briefing, (NICE, London) nice.org.uk/guidance/lgb10

PSHE Association, (2017), PSHE Education Programme of Study, Key Stages 1-5, PSHE, London

Ofsted, (2013), Not Good Enough yet: Personal, Social, Health and Economic Education in Schools, Ofsted, London (www.ofsted.gov.uk/resources/130065)

Schmidt, L., Garratt, A.M., and Fitzpatrick, R., (2001), Instruments for Children and Adolescents: A Review, Report form the Patient-reported Health Instruments Group, , (National Centre for Health Outcomes Development) to the UK Department of Health, Dept of Public Health, University of Oxford (http://phi.uhce.ox.ac.uk/)

Silk JS, Vanderbilt-Adriance E, Shaw DS, et al., (2007), "Resilience among children and adolescents at risk for depression: mediation and moderation across social and neurobiological contexts", Dev Psychopathol, 19: 841–65.

Thapar, A., Collishaw, S., Pine, D. et al, (2012), "Depression in Adolescence", Lancet, 379, 1058-1067

Weisz JR, McCarty CA, Valeri SM., (2006), "Effects of psychotherapy for depression in children and adolescents: a meta-analysis", Psychol Bull 2006; 132: 132–49:

Figure 1: Healthy Minds Program Flow

		Number of Lessons
Year 7 Managing the world around me	Penn Resilience Programme	18
	Media Navigator	6
	• breathe	6
Year 8 Moving toward my future	From School to Life	9
	Unplugged (Substance Use/Misuse) Part 1	4
	Media Influences	8
	Sex Ed Sorted Part 1	4
	Relationship Smarts	4
	Review and Connect	2
Year 9 Taking control of the decisions I make	Relationship Smarts (continued)	8
	School Health Alcohol Harm Reduction Project	6
	Sex Ed Sorted Part 2	6
	Resilient Decisions	4
Year 10 Preparing for the Future	Mental Illness Investigated	6
	Unplugged (Substance Use/Misuse) Part 2	8
	Parents Under Construction	10
	Resilient Learners	4

Notes: Figure 1 lists the lessons, all of which are based on evidence provided in Coleman et al (2011) and detail therein, in each of the four years that *Healthy Minds* runs in secondary schools.

Figure 2: Participant flow diagram

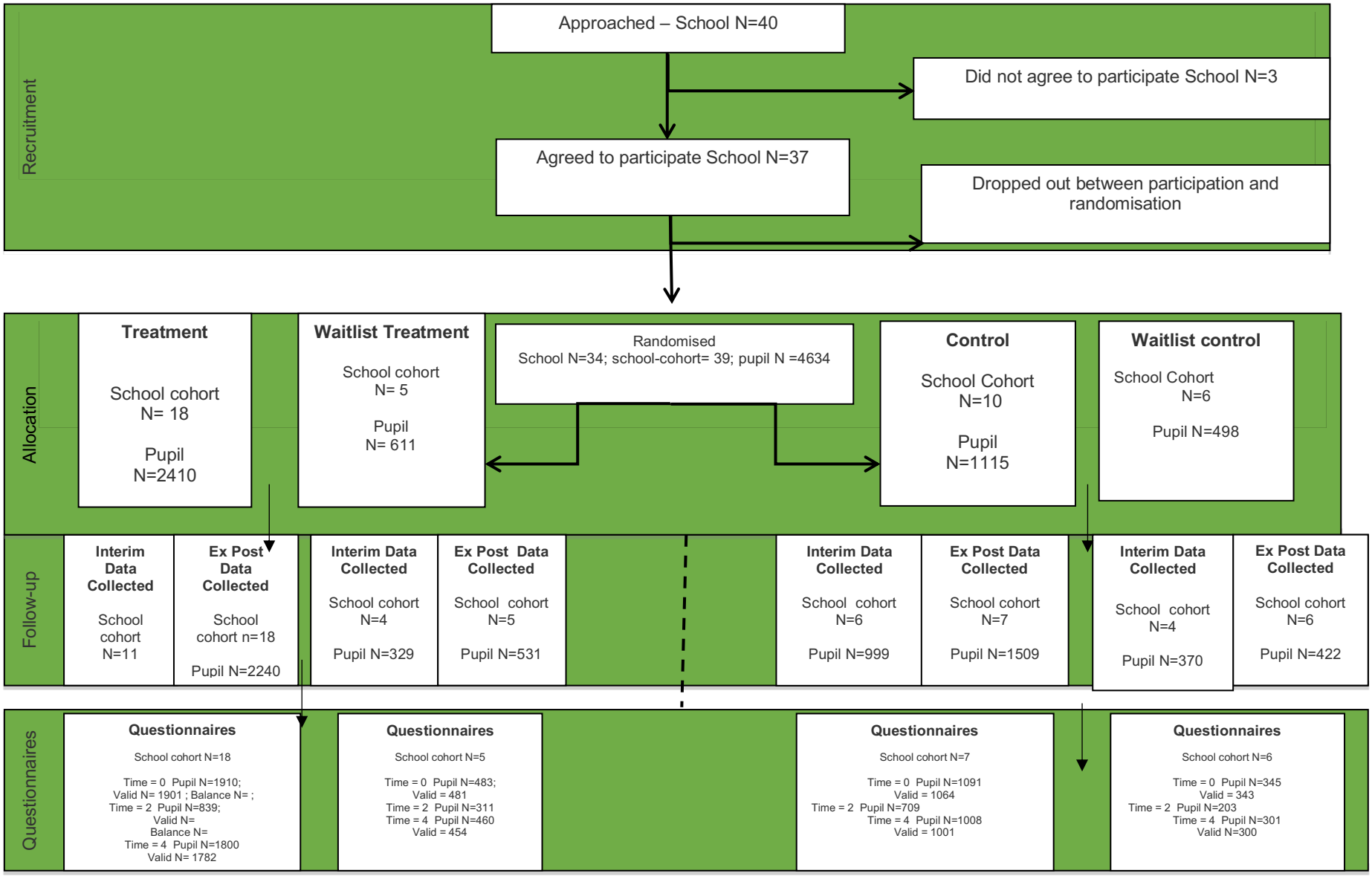


Figure 3: Effects of Reduced Items

Figure 3a) Intention to Treat Effects

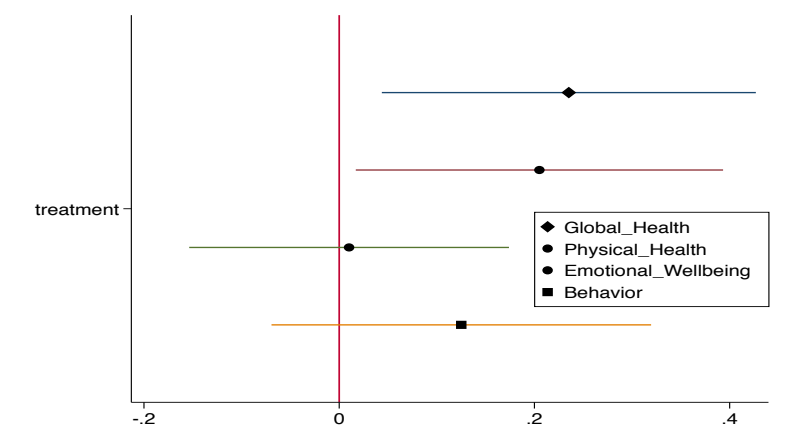


Figure 3b) Average Treatment Effect on Compliers

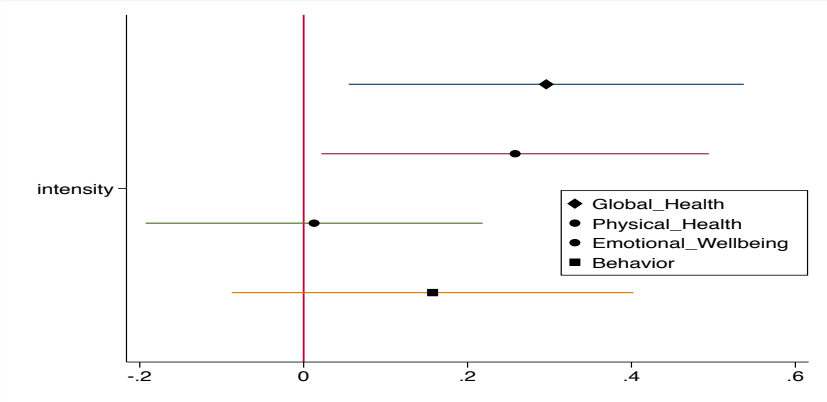


Figure 3c) IPW weights

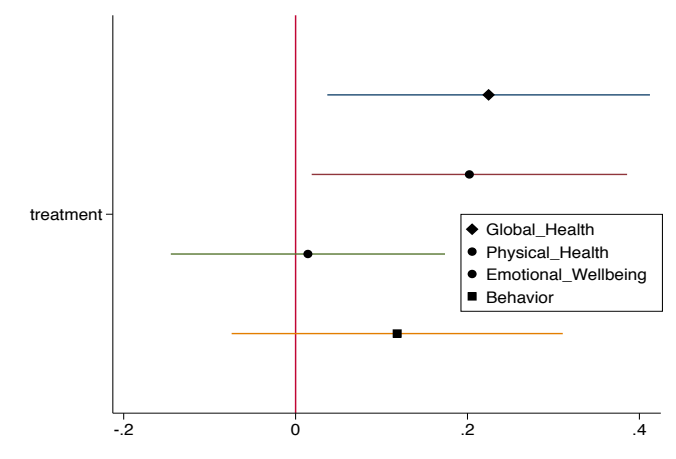
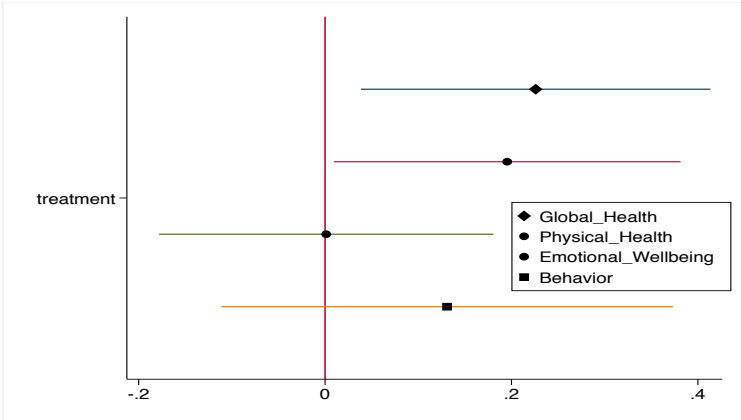


Figure 3d) Pupil Fixed Effect



Notes: See notes to Tables 5 through 7

Table 1: Study Timeline

Date	Activity
January 2013	**Approach potential participating schools**
January 2013 through September 2013	**Recruitment of schools**
April – September 2013	**Allocation to Phase 1 schools (pupils enter study September 2013) and Phase 2 schools (pupils enter study September 2014)**
July 2013	**Year 1 Teacher Training for Phase 1 schools**
September 2013	**DHMT teaching begins for Phase 1 schools**
September 2013	**Data (baseline) questionnaires administered to Phase 1 schools (Treatment and wait-list control schools)**
July 2014	**Year 1 Teacher Training for Phase 2 schools and Year 2 Teacher Training for Phase 1 schools**
September 2014	**DHMT teaching begins for Phase 2 schools**
September 2014	**Data (baseline) questionnaires administered to Phase 2 schools (Treatment, wait-list treatment schools and control schools)**
July 2015	**Year 2 Teacher Training for Phase 2 schools and Year 3 Teacher Training for Phase 1 schools**
May/June 2015	**Data (interim) questionnaires administered to Phase 1 schools (Treatment and wait-list control schools)**
July 2016	**Year 3 Teacher Training for Phase 2 schools and Year 4 Teacher Training for Phase 1 schools**
May/June 2016	**Data (interim) questionnaires administered to Phase 2 schools (Treatment, wait-list treatment and control schools)**
July 2017	**Year 4 Teacher Training for Phase 2 schools**
May/June 2017	**Data (endline) questionnaires administered to Phase 1 schools (Treatment and wait-list control schools)**
May/June 2018	**Data (endline) questionnaires administered to Phase 2 schools (Treatment, wait-list treatment and control schools)**
August/September 2018	**Data released and analysis undertaken**

Table 2: School cohorts by study stage

Number of School Cohorts in Each Study Stage			
Study Time	Time =0 (Baseline)	Time=2 (Interim)	Time =4 (Endline)
Phase 1 (2013)			
Treatment	7	5	7
Wait List Control	6	4	6
Phase 2 (2014)			
Wait List Treatment	5	4	4
Treatment	11	6	11
Control	10	6	7

Table 3: Questionnaire Scales and Item Number

CHQ 87 Scales	Item Number
Global Health	1
Physical Functioning	9
Emotional Difficulties	3
Behavioural Difficulties	3
Physical Difficulties	3
Self Esteem	14
Pain and Discomfort	2
Behaviour	16
Global Behaviour	1
Mental Health	16
General Health	11
Family Activities	6
Family Cohesion	1

Notes: Table 3 details the number of items (or questions) that make up each of the CHQ-87 scales that are the outcomes considered in this work.

Table 4: T tests of Differences at baseline (t=0)

Outcome	Pupil Level t=0	School Cohort Level t=0	Outcome	Pupil Level t=0	School Cohort Level t=0
Global Health	-0.009	-0.045	Mental Health	-0.104	0.075
	(0.036)	(0.034)		(0.037)	(0.113)
N	3789	39	N	3789	39
Physical Functioning	-0.082	0.038	Self Esteem	-0.078	-0.042
	(0.040)	(0.052)		(0.036)	(0.115)
N	3789	39	N	3789	39
Emotional Difficulties	-0.028	0.036	General Heath	-0.069	-0.021
	(0.039)	(0.097)		(0.036)	(0.065)
N	3789	39	N	3789	39
Behavioural Difficulties	-0.036	0.043	Family Activities	-0.093	-0.014
	(0.039)	(0.097)		(0.037)	(0.060)
N	3789	39	N	3789	39
Physical Difficulties	0.016	0.013	Family Cohesion	0.101	0.008
	(0.038)	(0.082)		(0.036)	(0.061)
N	3789	39	N	3789	39
Behaviour Difficulties	-0.162	-0.021	Pain and Discomfort	-0.032	-0.173
	(0.037)	(0.104)		(0.035)	(0.202)
N	3789	39	N	3789	39
Global Behaviour	-0.109	-0.039			
	(0.037)	(0.077)			
N	3789	39			

Table 5: Intention to Treat Analysis

	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
Treatment	0.235	0.122	0.049	0.057	-0.172
	(0.098)	(0.114)	(0.085)	(0.088)	(0.106)
B&H	0.019	0.738	1.000	1.000	0.150
N	7252	7252	7252	7252	7252
	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.288	0.239	0.127	0.044	0.001
	(0.106)	(0.108)	(0.109)	(0.096)	(0.096)
B&H	0.007	0.035	0.530	1.000	1.000
N	7252	7252	7252	7252	7252
	General Health	Family Activities	Family Cohesion		
Treatment	0.139	0.174	0.242		
	(0.104)	(0.102)	(0.101)		
B&H	0.332	0.219	0.001		
N	7252	7252	7252		

Notes: Models include school fixed effects. All outcomes are standardised to have a mean of 0 and a standard deviation of 1. Global health is the primary outcome of the trial. Standard errors are in brackets and are wild bootstrapped and clustered by student within school. B&H rows document the probability of significance after the Benjamini, Y. and Hochberg, Y. (1995) multiple comparison correction.

Table 6 Alternative Estimates

	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
Average Treatment Effect estimated by Two Stage Least Squares					
Treatment	0.296	0.153	0.061	0.072	-0.217
	(0.160)	(0.165)	(0.070)	(0.075)	(0.115)
B&H	0.092	1.000	1.000	0.878	0.078
N	7252	7252	7252	7252	7252
	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.363	0.300	0.160	0.001	0.055
	(0.140)	(0.148)	(0.121)	(0.111)	(0.075)
B&H	0.010	0.051	0.343	0.092	1.000
N	7252	7252	7252	7252	7252
	General Health	Family Activities	Family Cohesion		
Treatment	0.175	0.219	0.304		
	(0.107)	(0.184)	(0.113)		
B&H	0.164	0.504	0.007		
N	7252	7252	7252		
Estimates with Inverse Probability Weights					
	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
Treatment	0.212	0.109	0.029	0.053	-0.089
	(0.096)	(0.111)	(0.097)	(0.107)	(0.116)
B&H	0.029	0.708	1.000	1.000	1.000
N	7252	7252	7252	7252	7252
	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.311	0.206	0.155	0.040	-0.025
	(0.115)	(0.105)	(0.139)	(0.119)	(0.105)
B&H	0.007	0.065	0.492	1.000	1.000
N	7252	7252	7252	7252	7252
	General Health	Family Activities	Family Cohesion		
Treatment	0.150	0.177	0.245		
	(0.096)	(0.120)	(0.117)		
B&H	0.169	0.229	0.043		
N	7252	7252	7252		

Notes: Models include school fixed effects. All outcomes are standardised to have a mean of 0 and a standard deviation of 1. Global health is the primary outcome of the trial. Standard errors are in brackets and are wild bootstrapped and clustered by student within school. B&H rows document the probability of significance after the Benjamini, Y. and Hochberg, Y. (1995) multiple comparison correction.

Table 7 Pupil Fixed Effects

	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
Treatment	0.225	0.121	0.050	0.058	-0.156
	(0.096)	(0.111)	(0.083)	(0.086)	(0.103)
B&H	0.022	0.600	1.000	1.000	0.191
Estimator	RE	RE	RE	RE	RE
Hausman	p=0.563	p=0.124	p=0.642	p= 0.409	p=0.498
N	7252	7252	7252	7252	7252
	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.290	0.239	0.123	0.023	0.009
	(0.104)	(0.108)	(0.109)	(0.096)	(0.095)
B&H	0.005	0.035	0.475	1.000	1.000
Estimator	RE	RE	RE	RE	RE
Hausman	p= 0.650	p= 0.223	p=0.171	p= 0.103	p=0.555
N	7252	7252	7252	7252	7252
	General Health	Family Activities	Family Cohesion		
Treatment	0.134	-0.067	0.238		
	(0.101)	(0.116)	(0.101)		
B&H	0.299	0.899	0.020		
Estimator	RE	FE	RE		
Hausman	p=0.241	p=0.011	p=0.755		
N	7252	7252	7252		

Notes: Models include pupil fixed effects. All outcomes are standardised to have a mean of 0 and a standard deviation of 1. Global health is the primary outcome of the trial. Standard errors are in brackets and are wild bootstrapped and clustered by student within school. B&H rows document the probability of significance after the Benjamini, Y. and Hochberg, Y. (1995) multiple comparison correction.

Table 8: Gender Differences at age 11/12

Sample= Healthy Minds Study Sample								
Outcome	Male	Female	Outcome	Male	Female	Outcome	Male	Female
Internalising	0.059	0.033	Externalising	0.015	0.0341	Health	-0.059	0.042
	(0.958)	(0.870)		(0.937)	(0.960)		(0.913)	(0.794)
	1590	1872		1590	1872		1590	1872
Global Health	-0.010	-0.011	Physical Functioning	-0.137	-0.027	Emotional Difficulties	-0.002	-0.131
	(1.016)	(0.963)		(1.220)	(1.069)		(1.025)	(1.146)
N	1590	1872	N	1590	1872		1590	1872
Behavioural	-0.124	0.027	Self Esteem	0.030	0.005	Physical	-0.092	0.058
Difficulties	(1.130)	(0.989)		(0.995)	(0.980)	Difficulties	(1.119)	(0.915)
N	1590	1872	N	1590	1872	N	1590	1872
Pain and Discomfort	0.037	0.010	Behaviour	0.012	0.029	Global	0.009	0.035
	(0.992)	(1.005)		(1.017)	(1.003)	Behaviour	(0.983)	(1.005)
N	1590	1872	N	1590	1872	N	1590	1872
Mental Health	0.040	-0.002	General Health	0.035	0.041	Family	-0.019	0.012
	(1.033)	(0.993)		1.023	(1.003)	Activities	(0.997)	(1.001)
N	1590	1872		1590	1872	N	1590	1872
Family	-0.014	-0.012						
Cohesion	1.015	(0.981)						
N	1590	1872						
Sample= Millenium Cohort								
Conduct Problems	-0.089	0.090	Emotions	0.035	-0.036	Hyper Activity	-0.185	0.189
	(1.062)	(0.925)		(0.997)	(1.001)		(1.041)	(0.919)
N				6483	6483		1590	1872
Peer Problems	-0.055	0.057	Pro Social	-0.160	0.163			
	(1.046)	(0.947)		(1.077)	(0.885)			
	6483	6483		6483	6483			

Notes: In the bottom panel Data is from the Millennium Cohort Study which follows the lives of around 19,000 children born in the UK in 2000 and 2001. The data used is from parent responses on the SDQ questionnaire when the child was 11/12. See **http://www.sdqinfo.com/a0.html** for further details on the SDQ. All outcomes are back coded so higher scores indicate better outcomes.

Table 9: Difference in Difference Estimates with Gender Interactions

	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
Treatment	0.340	-0.214	-0.251	-0.206	-0.415
	(0.113)	(0.092)	(0.121)	(0.126)	(0.112)
B&H	0.005	0.037	0.101	0.328	0.000
Treatment*Male	-0.243	0.335	0.449	0.373	0.605
	(0.062)	(0.070)	(0.062)	(0.058)	(0.071)
B&H	0.000	0.000	0.000	0.000	0.000
Estimator	RE	FE	RE	RE	FE
Hausman test	0.083	0.020	0.543	0.487	0.000
N	5660	5660	5660	5660	5660

	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.202	0.453	-0.129	-0.153	-0.398
	(0.128)	(0.132)	(0.114)	(0.106)	(0.091)
B&H	0.494	0.000	1.000	0.001	0.000
Treatment*Male	0.116	-0.295	0.222	-0.096	0.607
	(0.060)	(0.054)	(0.065)	(0.056)	(0.060)
B&H	0.338	0.000	0.003	1.000	0.000
Estimator	RE	RE	RE	RE	RE
Hausman Test	0.319	0.325	0.407	0.135	0.864
N	5660	5660	5660	5660	5660

	General Health	Family Activities	Family Cohesion
Treatment	-0.128	-0.271	0.331
	(0.099)	(0.130)	(0.108)
B&H	1.000	0.078	0.000
Treatment*Male	0.323	0.223	-0.221
	(0.065)	(0.074)	(0.061)
B&H	0.000	0.013	0.000
Estimator	RE	FE	RE
Hausman Test	0.160	0.000	0.458
N	5660	5660	5660

Notes: see Notes to Table 5.

Table 10 Occupation Aspirations

Baseline Model: No Gender Interactions				
	Share of Males	Average Income	Average Hours	Non-Linear Hours
Treatment	-0.109	-2.700	-1.314	0.044
	(0.054)	(0.987)	(1.066)	(0.071)
B&H	0.067	0.006	0.872	1.000
N	2646	2646	2646	2646
	Competitiveness	People	Data	Things
Treatment	-0.347	0.283	-0.277	0.205
	(0.164)	(0.108)	(0.151)	(0.147)
B&H	0.005	0.009	0.136	0.434
N	2646	2646	2646	2646
Model with Gender Interactions				
	Share of Males	Average Income	Average Hours	Wage Hours Elasticity
Treatment	-0.001	-2.239	-0.132	-0.014
	(0.047)	(1.026)	(1.135)	(0.067)
B&H	0.075	0.029	0.907	0.834
Treatment*Male	-0.055	-0.482	-0.214	-0.007
	(0.031)	(0.879)	(0.715)	(0.034)
B&H	0.072	0.779	1.000	1.000
N	2646	2646	2646	2646
	Competitiveness	People	Data	Things
Treatment	-0.376	0.209	-0.336	0.396
	(0.214)	(0.175)	(0.190)	(0.174)
B&H	0.305	0.232	0.076	0.023
Treatment*Male	-0.135	-0.025	0.027	-0.043
	(0.131)	(0.119)	(0.146)	(0.127)
B&H	0.349	1.000	1.000	1.000
N	2646	2646	2646	2646

Notes: Models include school fixed effects. Standard errors are in brackets and are wild bootstrapped and clustered by student within school. B&H rows document the probability of significance after the Benjamini, Y. and Hochberg, Y. (1995) multiple comparison correction. Share of males lies between 0 and 1, average income and average hours are weekly, wage hours elasticity lies between -1.07 and 0.45 with higher values indicating less flexible work and competitiveness, people, data and things are standardised to have a mean of zero and a standard deviation of 1.

Table 9: Cost of delivering Healthy Minds

Item	Type of cost	Cost	Total cost over 4 years	Total cost per pupil per year over 4 years
One-off teacher training	*Teacher training cost per school*	£3,600 per school	£3,600 per school	
	Replacement teacher cost during training	£3,040 per school	£3,040 per school	
Total set-up cost per student per school year				£18.50 (assuming 1 teacher teaches 3 classes of 30 pupils each)
Material printing costs per student	*Material per student*	£5 per student school year	£600 (assuming 30 pupils per year)	£5
Total Cost			£7,250	£23.50

Appendix for:

Widening the High School Curriculum to Include Soft Skill Training: Impacts on Health, Behaviour, Emotional Wellbeing and Occupational Aspirations.

Not for publication

Grace Lordan LSE
Alistair McGuire LSE

Appendix A: Program Overview

Figure A.1 Details of the number of one-hour lessons attached to each module:

Notes: figure A.1 describes the flow of one-hour lessons in each of the four years that *Healthy Minds* runs in secondary schools.

Appendix B: Robustness analysis with no balance and imputations restrictions:
Table B.1: Intention to Treat Analysis

	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
Treatment	0.232	0.125	0.049	0.058	-0.168
	(0.108)	(0.089)	(0.074)	(0.095)	(0.102)
B&H	0.038	0.349	1.000	1.000	0.163
N	7326	7300	7278	7272	7290
	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.289	0.238	0.137	0.045	0.038
	(0.117)	(0.081)	(0.117)	(0.086)	(0.115)
B&H	0.015	0.003	0.632	1.000	1.000
N	7209	7212	7030	7051	6429
	General Health	Family Activities	Family Cohesion		
Treatment	0.143	0.184	0.244		
	(0.091)	(0.094)	(0.122)		
B&H	0.212	0.072	0.058		
N	7277	7177	7189		

Notes: Models include school fixed effects. All outcomes are standardised to have a mean of 0 and a standard deviation of 1. Global health is the primary outcome of the trial. Standard errors are in brackets and are wild bootstrapped and clustered by student within school. B&H rows document the probability of significance after the Benjamini, Y. and Hochberg, Y. (1995) multiple comparison correction.

Table B.2 Average Treatment Effect estimated by Two Stage Least Squares

	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
Treatment	0.292	0.157	0.062	0.073	-0.211
	(0.137)	(0.112)	(0.093)	(0.119)	(0.128)
B&H	0.039	0.349	1.000	1.000	0.161
N	7326	7300	7278	7272	7290
	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.363	0.299	0.172	0.047	0.056
	(0.149)	(0.102)	(0.147)	(0.144)	(0.108)
B&H	0.016	0.003	0.629	1.000	1.000
N	7209	7212	7030	6429	7051
	General Health	Family Activities	Family Cohesion		
Treatment	0.180	0.231	0.307		
	(0.115)	(0.118)	(0.153)		
B&H	0.217	0.072	0.059		
N	7277	7177	7189		

Notes: Models include school fixed effects. All outcomes are standardised to have a mean of 0 and a standard deviation of 1. Global health is the primary outcome of the trial. Standard errors are in brackets and are wild bootstrapped and clustered by student within school. B&H rows document the probability of significance after the Benjamini, Y. and Hochberg, Y. (1995) multiple comparison correction.

Table B.3 Intention to Treat Model with Inverse Probability Weights

	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
Treatment	0.236	0.081	0.046	0.063	-0.050
	(0.100)	(0.114)	(0.096)	(0.105)	(0.119)
B&H	0.018	0.878	1.000	1.000	1.000
N	6581	6573	6557	6554	6564
	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.299	0.189	0.087	-0.057	0.008
	(0.109)	(0.115)	(0.144)	(0.123)	(0.112)
B&H	0.006	0.119	1.000	1.000	1.000
N	6527	6530	6461	6471	6141
	General Health	Family Activities	Family Cohesion		
Treatment	0.162	0.143	0.182		
	(0.102)	(0.127)	(0.116)		
B&H	0.145	0.424	0.169		
N	6554	6520	6517		

Notes: Models include school fixed effects. All outcomes are standardised to have a mean of 0 and a standard deviation of 1. Global health is the primary outcome of the trial. Standard errors are in brackets and are wild bootstrapped and clustered by student within school. B&H rows document the probability of significance after the Benjamini, Y. and Hochberg, Y. (1995) multiple comparison correction. Numbers are smaller than Tables B.1 and B.2 because pupils who join in final phase are excluded from the analysis

Table B.4 Intention to Treat model with Pupil Fixed Effects

	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
			Pupil Fixed Effects		
Treatment	0.223	0.124	0.051	0.058	-0.151
	(0.108)	(0.088)	(0.073)	(0.093)	(0.101)
B&H	0.022	0.600	1.000	1.000	0.191
Estimator	RE	RE	RE	RE	RE
Hausman	p=0.570	p=0.156	p=0.635	p=0.407	p=0.498
N	7326	7300	7278	7272	7290
	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.291	0.238	0.131	0.022	0.045
	(0.115)	(0.081)	(0.116)	(0.086)	(0.114)
B&H	0.005	0.035	0.475	1.000	1.000
Estimator	RE	RE	RE	RE	RE
Hausman	p= 0.672	p=0.222	p=0.081	p= 0.068	p=0.640
N	7209	7212	7030	7051	6429
	General Health	Family Activities	Family Cohesion		
Treatment	0.138	-0.071	0.240		
	(0.089)	(0.141)	(0.121)		
B&H	0.299	0.899	0.020		
Estimator	RE	FE	RE		
Hausman	p=0.221	p=0.007	p=0.788		
N	7277	7177	7189		

Notes: Models include pupil fixed effects. All outcomes are standardised to have a mean of 0 and a standard deviation of 1. Global health is the primary outcome of the trial. Standard errors are in brackets and are wild bootstrapped and clustered by student within school. B&H rows document the probability of significance after the Benjamini, Y. and Hochberg, Y. (1995) multiple comparison correction.

Table B.5 Intention to Treat model with Pupil Fixed Effects and Gender Interactions

	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
Treatment	0.345	-0.193	-0.245	-0.202	-0.405
	(0.123)	(0.083)	(0.123)	(0.133)	(0.122)
B&H	0.007	0.032	0.085	0.338	0.001
Treatment*Male	-0.248	0.297	0.443	0.372	0.591
	(0.046)	(0.066)	(0.073)	(0.072)	(0.057)
B&H	0.000	0.000	0.000	0.000	0.000
Estimator	RE	FE	RE	RE	FE
Hausman test	0.066	0.030	0.512	0.458	0.000
N	5660	5660	5660	5660	5660

	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.207	0.445	-0.121	-0.142	-0.396
	(0.145)	(0.107)	(0.097)	(0.125)	(0.122)
B&H	0.494	0.000	0.932	1.000	0.001
Treatment*Male	0.111	-0.292	0.226	-0.099	0.618
	(0.063)	(0.049)	(0.061)	(0.058)	(0.063)
B&H	0.500	0.000	0.000	1.000	0.000
Estimator	RE	RE	RE	RE	RE
Hausman Test	0.393	0.361	0.224	0.062	0.796
N	5660	5660	5660	5660	5660

	General Health	Family Activities	Family Cohesion
Treatment	-0.129	-0.274	0.339
	(0.118)	(0.129)	(0.102)
B&H	1.000	0.169	0.001
Treatment*Male	0.325	0.229	-0.221
	(0.055)	(0.082)	(0.053)
B&H	0.000	0.022	0.000
Estimator	RE	FE	RE
Hausman Test	0.100	0.000	0.398
N	5660	5660	5660

Notes: Models include pupil fixed effects. All outcomes are standardised to have a mean of 0 and a standard deviation of 1. Global health is the primary outcome of the trial. Standard errors are in brackets and are wild bootstrapped and clustered by student within school. B&H rows document the probability of significance after the Benjamini, Y. and Hochberg, Y. (1995) multiple comparison correction. Numbers are smaller than table B.1 and B.2 because respondents needed to supply their gender information which was only asked in the final phase of the study and the end of the questionnaire. Therefore, i) pupils who did not respond in the final phase are excluded and ii) pupils who did not supply gender information are excluded.

Appendix C: Exploratory Factor Analysis of HRQoL Domains

We reduce the dimensionality of twelve CHQ_87 outcomes (we exclude the primary outcome from the dimensionality reduction), into three aggregate variables that combine outcomes related to physical health, emotional wellbeing and behaviour respectively. To achieve this we choose the items in the CHQ-CF87 secondary outcomes, which relate to these three distinct aggregate domains. Specifically, physical health comprises of the individual domains of physical health, family activities, general health and physical functioning. Emotional health comprises of self-esteem, mental health, and emotional difficulties. Finally, behaviour comprises of behavioural difficulties, global behaviour, family activities and behaviour. We note that family activities is included in health and behavior factors because the question specifically relates to 'how often has your health and behaviour' negatively affected family activities in six distinct ways.

We combine these three subsets of using exploratory factor analysis. Confirmatory factor analysis (CFA) is performed to extract the final latent variables. (Gorsuch, 2003; Thomson, 2004). Each factor is separately created by confirmatory factor analysis on the first rotation of the exploratory factor analysis described in the methods. The loadings on each latent factor created are as follows:

1. Physical health: physical health loading = 0.650, family activities loading =0.477, general health loading = 0.446 and physical functioning loading = 0.689.
2. Emotional health= self-esteem loading = 0.783, mental health loading=0.893, and emotional difficulties=0.500.
3. Behavior= behavioural difficulties loading=0.533, global behaviour loading=0.542, family activities loading = 0.591 and behaviour loading=0.868.

Appendix D: Robustness analysis with Sample Restricted to Pupils having supplied Gender Information:

Table D.1 Intention to treat analysis

	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
Treatment	0.249	0.131	-0.058	-0.043	-0.325
	(0.108)	(0.114)	(0.118)	(0.123)	(0.084)
B&H	0.025	0.544	1.000	1.000	0.000
N	5666	5666	5666	5666	5666
	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.259	0.328	-0.031	0.173	-0.156
	(0.128)	(0.125)	(0.111)	(0.104)	(0.093)
B&H	0.017	0.010	1.000	0.153	0.176
N	5666	5666	5666	5666	5666
	General Health	Family Activities	Family Cohesion		
Treatment	0.013	0.030	0.239		
	(0.097)	(0.116)	(0.105)		
B&H	1.000	1.000	0.030		
N	5666	5666	5666		

Notes: Models include school fixed effects. All outcomes are standardised to have a mean of 0 and a standard deviation of 1. Global health is the primary outcome of the trial. Standard errors are in brackets and are wild bootstrapped and clustered by student within school. B&H rows document the probability of significance after the Benjamini, Y. and Hochberg, Y. (1995) multiple comparison correction. Numbers are smaller than Tables B.1 and B.2 because pupils who join in final phase are excluded from the analysis

Table D.2 Average Treatment Effect Estimated by Two Stage Least Squares

	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
Treatment	0.317	0.173	-0.072	-0.051	-0.413
	(0.117)	(0.136)	(0.140)	(0.139)	(0.154)
B&H	0.008	0.379	1.000	1.000	0.007
N	5666	5666	5666	5666	5666
	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.337	0.420	-0.040	-0.197	-0.220
	(0.134)	(0.170)	(0.160)	(0.189)	(0.138)
B&H	0.014	0.017	1.000	0.182	0.644
N	5666	5666	5666	5666	5666
	General Health	Family Activities	Family Cohesion		
Treatment	0.017	0.038	0.306		
	(0.125)	(0.132)	(0.142)		
B&H	1.000	0.072	0.046		
N	5666	5666	5666		

Notes: Models include school fixed effects. All outcomes are standardised to have a mean of 0 and a standard deviation of 1. Global health is the primary outcome of the trial. Standard errors are in brackets and are wild bootstrapped and clustered by student within school. B&H rows document the probability of significance after the Benjamini, Y. and Hochberg, Y. (1995) multiple comparison correction. Numbers are smaller than Tables B.1 and B.2 because pupils who join in final phase are excluded from the analysis

Table D.3 Intention to Treat Model with Inverse Probability Weights:

	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
Treatment	0.237	0.060	-0.083	-0.075	-0.284
	(0.106)	(0.104)	(0.106)	(0.112)	(0.103)
B&H	0.034	1.000	0.936	1.000	0.006
N	5666	5666	5666	5666	5666
	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.218	0.308	-0.028	-0.211	-0.169
	(0.113)	(0.122)	(0.120)	(0.116)	(0.114)
B&H	0.008	0.012	1.000	0.113	0.254
N	5666	5666	5666	5666	5666
	General Health	Family Activities	Family Cohesion		
Treatment	0.018	0.001	0.284		
	(0.101)	(0.112)	(0.112)		
B&H	1.000	1.000	0.014		
N	5666	5666	5666		

Notes: Models include school fixed effects. All outcomes are standardised to have a mean of 0 and a standard deviation of 1. Global health is the primary outcome of the trial. Standard errors are in brackets and are wild bootstrapped and clustered by student within school. B&H rows document the probability of significance after the Benjamini, Y. and Hochberg, Y. (1995) multiple comparison correction. Numbers are smaller than Tables B.1 and B.2 because pupils who join in final phase are excluded from the analysis

Table D.4 Intention to Treat model with Pupil Fixed Effects

	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
Treatment	0.238	-0.052	-0.057	-0.043	-0.301
	(0.091)	(0.106)	(0.108)	(0.107)	(0.118)
B&H	0.009	1.000	1.000	1.000	0.012
Estimator	RE	FE	RE	RE	RE
Hausman	p=0.357	p=0.024	p=0.619	p=0.616	p=0.072
N	5666	5666	5666	5666	5666
	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.258	0.328	-0.033	0.191	-0.142
	(0.103)	(0.133)	(0.120)	(0.107)	(0.145)
B&H	0.014	0.018	1.000	0.122	0.613
Estimator	RE	RE	RE	RE	RE
Hausman	p= 0.521	p=0.070	p=0.459	p= 0.298	p=0.555
N	5666	5666	5666	5666	5666
	General Health	Family Activities	Family Cohesion		
Treatment	0.009	0.015	0.239		
	(0.097)	(0.101)	(0.111)		
B&H	1.000	0.828	0.045		
Estimator	RE	FE	RE		
Hausman	p=0.626	p=0.018	p=0.428		
N	5666	5666	5666		

Notes: Models include pupil fixed effects. All outcomes are standardised to have a mean of 0 and a standard deviation of 1. Global health is the primary outcome of the trial. Standard errors are in brackets and are wild bootstrapped and clustered by student within school. B&H rows document the probability of significance after the Benjamini, Y. and Hochberg, Y. (1995) multiple comparison correction.

Appendix E:

Aspirations Outcomes:

Our occupations averages are owed to Lekfuangfu and Lordan (2018). Specifically, the occupation averages draw on the Quarterly Labour Force Survey (QLFS). The QLFS is the main survey of individual economic activity in the Britain, and provides the official measure of the national unemployment rate. From this data occupation averages - for income, hours and share of male employment - were calculated by the authors at a three-digit SOC00 occupation level. We match this data directly to the pupil's aspirational occupational codes.

We also rely on a variable to proxy the wage-hours elasticity defined by Goldin (2014). Goldin interprets this occupation specific elasticity as capturing the wage penalty arising from working shorter hours: high elasticities imply a penalty for workers seeking short hours and indicate a lack of flexibility. Specifically, this variable was created by running a regression of the log of wages on log hours, occupation fixed effects, the interaction between log hours and the occupation fixed effects and a number of other controls[30] using the QLFS and consistent three-digit SOC00 codes. The proxy is then the coefficients on the interaction between occupation and log hours.

Our analysis also utlises three variables which capture what the job is about. These variables were created following the approach described by Lordan and Pischke (2016). Specifically, the definitions are based on O*NET version 5 items relating to the activities and context of an individual's work. These items on activities and context are linked to US Standard Occupation Codes (SOC) 2000. These 79 items report the level at which an occupation has a particular characteristic from 1 to 7. For example, in activities, an item might describe to which degree an occupation involves 'assisting and caring for others,' 'analyzing data or information,' or the 'repairing and maintaining of mechanical equipment.' Examples for context are the level of 'contact with others,' 'the importance of being exact or accurate,' and 'being exposed to hazardous conditions.' US SOC00 codes in the O*NET data were matched to the British SOC00. The O*NET items were then matched to the QLFS using the British SOC00 codes. Three latent factors '*people*,' '*brains*,' and '*brawn*' (PBB) are calculated using this data. We match these three factors for each occupation to pupil's aspired to occupation using the British SOC00 codes.

[30] The controls follow Goldin (2014). These are gender, age, age squared, age to the power of three, age to the power of four, education, ethnicity and year dummies.

The measure of occupation competitiveness is drawn from O*NET database version 15. Specifically, incumbents are asked: *"To what extent does this job require the worker to compete or to be aware of competitive pressures?"* with response options of 'not at all competitive' 'slightly competitive' 'moderately competitive' 'highly competitive' and 'extremely competitive'. This variable is standardised to have a mean of zero and a standard deviation of 1, and matched to the pupil's occupational aspirations in the same manner described for the PBB factors.

Appendix F: Robustness to Clustering Choice

The table below reproduces the main results from this work (Table 5 in main text) with standard errors clustered at the school level.

Table F.1: Intention to Treat Analysis

	Global Health	Physical Functioning	Emotional Difficulties	Behavioural Difficulties	Self Esteem
Treatment	0.235	0.122	0.049	0.057	-0.172
	(0.121)	(0.111)	(0.050)	(0.047)	(0.061)
B&H	0.104	0.919	1.000	0.608	0.011
N	7252	7252	7252	7252	7252
	Physical Difficulties	Pain and Discomfort	Behaviour	Global Behaviour	Mental Health
Treatment	0.288	0.239	0.127	0.044	0.001
	(0.080)	(0.055)	(0.057)	(0.055)	(0.073)
B&H	0.001	0.000	0.052	1.000	1.000
N	7252	7252	7252	7252	7252
	General Health	Family Activities	Family Cohesion		
Treatment	0.139	0.174	0.242		
	(0.064)	(0.097)	(0.067)		
B&H	0.128	0.184	0.001		
N	7252	7252	7252		

Notes: Models include school fixed effects. All outcomes are standardised to have a mean of 0 and a standard deviation of 1. Global health is the primary outcome of the trial. Standard errors are clustered by school. B&H rows document the probability of significance after the Benjamini, Y. and Hochberg, Y. (1995) multiple comparison correction.