IFS

Institute for Fiscal Studies

Emilia Del Bono
Ben Etheridge
Paul Garcia

Working paper

26/42

The economic value of childhood socio-economic skills

UKRI Economic and Social Research Council

The Economic Value of Childhood Socio-Emotional Skills[*]

Emilia Del Bono,[†] Ben Etheridge,[‡] Paul Garcia[§]

30th June 2026

Abstract

We examine the relationship between childhood socio-emotional skills and labour market outcomes using the 1970 British Cohort Study. A new factorization of teacher assessments of child behaviours reveals four skill dimensions, captured by attention, conduct, emotional, and peer problems. We find that conduct problems are linked to positive labour market outcomes, including higher earnings, increased labour supply and greater job satisfaction, with no significant effect on schooling. Conversely, attention problems correlate negatively with labour outcomes, partly due to educational effects. Our findings suggest the importance of early interventions which carefully distinguish between different socio-emotional problems.

JEL Classification: J24, J62, I21.

Keywords: Socio-emotional Skills; Human Capital; Child Development; Labour Supply; Occupational Sorting.

[*]Thanks to Uta Bolt, Massimiliano Bratti, Gabriella Conti, Eric French, Nina Guyon, Angus Holford, Michelle Jin, Giorgia Menta, Reinhard Pekrun, Birgitta Rabe, Sunčica Vujić, Andrea Weber, Rudolf Winter-Ebmer, and Yu Zheng for helpful discussions. We are grateful for comments and suggestions at the following conferences: 2023 ESPE (Belgrade), 2023 EALE (Prague), 2024 AIEL (Naples), 2024 IZA/ECONtribute (Bonn), 2024 Austrian Labour Economics Workshop as well as at seminars in Essex, Milan University (DEMM), LISER, IFS and Erasmus University Rotterdam. We also thank the editor and three anonymous referees for helpful suggestions. All mistakes and omissions are the authors' responsibility. This work was funded by the ESRC Research Centre on Micro-Social Change (MiSoC), award number ES/S012486/1.

[†]University of Essex and IZA. E-mail: edelbono@essex.ac.uk

[‡]University of Essex and IFS. E-mail: bsethe@essex.ac.uk

[§]University of Essex. E-mail: pagarcia@essex.ac.uk

1 Introduction

Advances in automation and artificial intelligence are changing the nature and composition of work, as many routine and even some non-routine tasks can now be performed by advanced technologies (Acemoglu & Restrepo, 2018, 2019; Autor, 2022). These developments have renewed interest in characterizing the human abilities that complement such technologies, particularly socio-emotional skills such as communication, leadership, and cooperation (Deming, 2017; Edin et al., 2022; Izadi & Tuhkuri, 2024). Understanding the economic value of these abilities requires tracing their origins in childhood, where the foundations of socio-emotional development are laid and where interventions are most likely to be effective.

To examine the economic value of childhood socio-emotional skills, a challenge is to establish how these skills should be defined and measured. Broadly, socio-emotional skills refer to the capacity to regulate emotions and behaviours and to interact effectively with others, forming part of a wider set of non-cognitive abilities that include goals, motivation, and preferences (Borghans et al., 2008; Kautz et al., 2014). Childhood measures of these skills predict a range of adult outcomes (Borghans et al., 2014; Goodman et al., 2015) and these skills are malleable to intervention (Alan et al., 2021; Sorrenti et al., 2025). Yet the literature has not reached a consensus on how best to represent them, making it difficult to determine which aspects matter most for labour market outcomes and where policy intervention is most effective.

In this paper we use data from the 1970 British Cohort Study (BCS70), a longitudinal survey following a cohort from birth into adulthood, to examine how variation in early socio-emotional skills relates to later economic outcomes. We measure skills at age 10 and consider their association with schooling, earnings, hours of work, occupational sorting, and other adult outcomes between the ages of 26 and 46. We assess robustness to an extensive set of confounders, explore mediating pathways through schooling as well as information gathered at age 16 – socialization, competitive activities, career interests, and mental health – and present new evidence on the role of these skills in the intergenerational transmission of socio-economic (SES) status.

Our first contribution is a new factorization of childhood socio-emotional skills using teacher-reported items from several psychological instruments – the Rutter Child Behaviour Questionnaire (Rutter, 1967), the Conners Hyperactivity Rating Scale (Conners, 1969), and the Swansea Assessment Battery (Butler et al., 1980). Using exploratory factor analysis, we arrive at a four-factor representation capturing 'attention' problems (e.g. difficulty concentrating or completing a task), 'conduct' problems (e.g. aggression, impulsivity), 'emotional' problems (e.g. anxiety, fear of new situations), and 'peer' problems (e.g. shyness, difficulty forming friendships). This differs from recent studies, which use a more restricted set of items and typically work with broader constructs such as 'internalising' and 'externalising' behaviours (e.g., Attanasio, De Paula, & Toppeta, 2020; Papageorge et al., 2019).

We then offer a comprehensive analysis of this representation. We show that our four-factor model maps into the main sub-scales of the Strengths and Difficulties Questionnaire (SDQ), one of the most widely used screening tools for behavioural and emotional difficulties in children (Goodman, 1997). The SDQ is now routinely collected in many social and economic longitudinal surveys, including the US Panel Study of Income Dynamics (PSID) and the UK Millennium Cohort Study (MCS).[1] Although related to measures such as the 'Big-Five' personality traits (Lewis et al., 2014), the SDQ is designed for children aged 3–17, accommodates both teacher and parent assessments, and directly captures behavioural dimensions that are the focus of educational and clinical interventions (Goodman et al., 2000; Hall et al., 2019). Personality measures, by contrast, were developed for adults and rely on self-reports, which are less reliable in children (Vicentini et al., 2025). Our approach therefore bridges the personality and child-development literatures, while being of more direct policy relevance.

Our second contribution is that this finer representation – in particular, distinguishing conduct from attention within the broader externalising construct – reveals patterns obscured in existing work. Whereas attention problems predict substantially poorer education and labour market outcomes, conduct problems are *positively* associated with earnings. The conduct-earnings association is meaningful: a one standard deviation increase in conduct problems is associated with earnings around 3½ percent higher.[2] Attention, emotion, and peer problems are all predictive of negative outcomes, with attention deficits the most important quantitatively. The positive conduct-earnings relationship reflects higher wages as well as longer working hours and is robust to a wide range of controls for family background, neighbourhood, and school environment. Our results also replicate in the NLSY79 Child and Young Adult survey (NLSY79 CYA), which allows us to net out unobserved family effects through sibling comparisons, reducing concerns about omitted variables.

This distinction between conduct and attention echoes recent findings. Papageorge et al. (2019) report that externalising behaviours – which we show conflate attention and conduct – predict better labour market outcomes despite lower schooling attainment. We clarify that these apparently contrasting results for schooling and earnings reflect different underlying skills: attention versus conduct.[3] As we argue below, this distinction matters when considering what schools can do to improve child outcomes.

Our third contribution is new evidence on the pathways mediating the relationship between early skills and later outcomes. Beyond education, we consider participation in sport competition,

[1]The SDQ has also been included in the Avon Longitudinal Study of Parents and Children, the UK Understanding Society, the German Socio-Economic Panel, several Danish cohort studies (Niclasen et al., 2012), and the Longitudinal Study of Australian Children.

[2]By comparison, the effect of cognition is a little over 9 percent.

[3]A related conclusion is drawn in the comprehensive review by Nicoletti and Vidiella-Martin (2025), who survey the economics and epidemiology literature on ADHD. They report consistent quasi-experimental evidence showing that inattention, and not features related to conduct problems, is the main driver of adverse education outcomes.

socialization, career interests, and mental health, all measured around age 16. These age-16 mediators are strongly related to age-10 skills: attention problems predict lower interest in a business career, emotional problems predict worse adolescent mental health, and, most notably, conduct problems predict participation in sport competition – suggesting that children with conduct problems may also be highly competitive.[4] Yet *none* of these mediators explains the link between early skills and adult labour market outcomes, suggesting that childhood skills capture adult competencies that our adolescent measures only partially reflect. Educational attainment itself partially mediates the effect of attention problems but does not capture the role of the other age-10 skills; if anything, we show that years of schooling *compensate* for emotional and peer problems.

We additionally consider the contribution of childhood socio-emotional skills to SES inequalities in adulthood, using the Kitagawa-Oaxaca-Blinder decomposition (Fortin et al., 2011). We show that SES gaps in adult outcomes are driven mainly by differences in skills endowments rather than in returns to skills. Once again, conduct and attention play distinct roles: differences in conduct reduce SES gaps, while differences in attention widen them. Here, however, differences in cognition account for the bulk of the explained gap. This is consistent with other studies that have analysed the role of cognitive and non-cognitive skills in the transmission of economic (dis)advantage. Our contribution is to use a formal latent-factor model rather than aggregating childhood behaviours into pre-defined indices (as in Blanden et al., 2007), and analyse the distinct roles of different socio-emotional skills rather than a single non-cognitive factor (Bolt et al., 2021).

Beyond the labour market, our skills measures also predict a broad range of adult health and well-being outcomes. Most strikingly, conduct problems carry a mixed profile: higher job satisfaction alongside higher rates of smoking, drinking, and arrests, while problems with emotions and relating to peers predict worse mental health and lower life satisfaction.

Taken together, these findings speak to several policy issues. First, they support interventions focused on the early years: childhood skills retain their long-term economic value even after accounting for intermediate outcomes and mechanisms. Second, if the aim is to raise educational attainment and narrow SES earnings gaps, interventions should focus on attention problems; conduct, emotional and peer problems appear less relevant in this respect. Third, the positive association between conduct problems and labour market outcomes suggests that what is sometimes read as misbehaviour may be the manifestation of skills, such as assertiveness or competitiveness, that children find difficult to regulate at a young age. This suggests a less punitive response to classroom misbehaviour may be suitable.

This positive labour-market association does not mean that conduct problems are uniformly beneficial: as just noted, they predict more risky behaviours in adulthood, and are correlated with emotional problems and negatively with cognition, both of which themselves predict poorer

[4]This interpretation is also supported by studies that link competitive, or aggressive behaviour with success in entrepreneurial or high-stakes careers (see for example Levine & Rubinstein, 2017).

labour-market outcomes. The appropriate response is therefore not to encourage misbehaviour, but to help children regulate and redirect these tendencies. This aligns with the movement in education advocating restorative justice practices in schools (Darling-Hammond & Fronius, 2022; Hopkins, 2023; Zakszeski & Rutherford, 2021), which recent evidence suggests can reduce both in-school suspensions and out-of-school arrests (Adukia et al., 2025).

After a review of the literature, Section 3 describes the data and Section 4 the methodology. Section 5 presents the factor analysis, Section 6 examines long-run outcomes, and Section 7 explores intergenerational SES transmission. Section 8 concludes.

2 Related literature

Our paper relates to two principal strands of the literature. The first examines the *long-term returns to non-cognitive or socio-emotional skills measured in childhood*, typically from mother-reported behaviour mapped onto personality traits (see Almlund et al., 2011; Borghans et al., 2008).[5] Within this literature, findings on which specific behaviours matter are mixed. Using the same BCS70 data we analyse here, Prevoo and ter Weel (2015) find that conscientiousness, closely (and inversely) related to our attention factor, is positively associated with labour market outcomes, but report no association for agreeableness, closely (and inversely) related to our conduct factor. Goodman et al. (2015), also using BCS70, derive a similar four-dimensional representation to ours from mother-reported Rutter items but find that *good* conduct predicts higher educational attainment and earnings, the opposite of our result.[6] Attanasio, De Paula, and Toppeta (2020) use a two-factor representation derived from selected BCS70 maternal Rutter items and find that externalising problems in particular are strongly and negatively associated with adult employment and earnings.

Our analysis departs from these studies in three key respects. First, we provide a more comprehensive representation of childhood socio-emotional skills. Drawing on a broader set of behavioural items than is typical, we identify four latent factors that distinguish attention deficits from conduct problems and align closely with the sub-scales of the Strengths and Difficulties Questionnaire, one of the most widely used instruments in research on child development (e.g., Fitzsimons & Vera-Hernández, 2022). Second, our skills measures are derived using dedicated factor analysis, which allows us to address issues of measurement error using the approach in Heckman et al. (2013) and Bolt et al. (2021). Third, we rely on teacher reports and show that they are more informative in terms of labour market predictions than mother reports, usually preferred in these

[5]The seminal work of Heckman et al. (2006) analyses the effect on later outcomes of youth-measured skills aggregated into cognitive and non-cognitive dimensions, with the non-cognitive factor combining many of the elements analysed in further detail in the subsequent literature.

[6]Using the National Education Longitudinal Survey, Segal (2013) shows that teacher-reported 'misbehaviour' predicts lower earnings overall, but further disaggregation reveals that disruptive behaviour—closest to our conduct factor—carries a positive earnings association.

earlier studies.[7] In fact, we offer an explicit multitrait-multimethod analysis of mother reports alongside our main measures from teachers, as in Goodman et al. (2010). In so doing we contribute to a recent literature that examines the measurements of skills from different informants (e.g., Del Bono et al., 2026; Feng et al., 2022; Johnston et al., 2014).[8]

A second strand examines the *labour market returns to social skills measured in adolescence or early adulthood*. Weinberger (2014) shows that social skills proxied by high-school sports participation and leadership roles predict wages and occupational attainment, with returns rising over time, particularly for individuals with strong cognitive skills. Deming (2017) documents a similar rise using NLSY data and finds parallel growth in employment in jobs intensive in interpersonal tasks.[9] A related set of studies exploits mandatory military assessments: Edin et al. (2022) use Swedish draft records to construct a non-cognitive index combining social maturity, leadership, and emotional stability, and show its return rose sharply between 1992 and 2013; Izadi and Tuhkuri (2024) use Finnish conscription data and attribute rising returns primarily to extraversion.[10]

We speak to this literature from a child development perspective. Our skill definitions are derived from validated psychological batteries rather than proxies such as sport or club participation (as in e.g. Deming, 2017), and our measures are observed at age 10 rather than in early adulthood (as in e.g. Izadi & Tuhkuri, 2024), so they precede educational and labour market decisions. This early measurement also lets us explore the mediating pathways, including educational attainment, but also competitive sport, socialization, career interests, and adolescent mental health, through which childhood behaviours may translate into adult outcomes.

Finally, our evidence of a positive relationship between conduct problems and labour market success connects to work in evolutionary psychology interpreting aggressive behaviour as context-dependent adaptation: advantageous in competition for status or partners, but detrimental in other settings (Buss & Shackelford, 1997; Volk et al., 2022). Consistent with this view, we also find that individuals scoring high on conduct problems in childhood are more likely to drink, smoke, and engage in criminal behaviour as adults.

[7] Papageorge et al. (2019) primarily use teacher-reported measures of child socio-emotional skills from the National Child Development Study (NCDS), but also assess teacher vs mother reports using the PSID. Conti et al. (2014) use *all* available measures of child behaviour in the BCS70 at age 10 - whether mother reported, teacher reported, or self-reported - and identify different latent factors, with some factors being mainly driven by maternal reports and others by teacher reports.

[8] The idea of relying on multiple evaluators or multiple measurement methods has a long history in the broader psychometric and applied statistics fields (Campbell & Fiske, 1959; Joreskog, 1971).

[9] Similar findings for the UK are documented by Dickerson and Morris (2019) and, for an earlier period, by Borghans et al. (2014).

[10] Using data from Finnish military assessments, Jokela et al. (2017) provide evidence of the *supply* of skills by documenting an increase in the supply of self-confidence, sociability, and leadership motivation across cohorts born in the 1960s and 1970s.

3 Data

3.1 BCS70

Our primary data source is the 1970 British Cohort Study (BCS70), which follows approximately 17,000 individuals born in England, Scotland, and Wales during the second week of April 1970, with survey waves conducted at ages 5, 10, 16, 26, 30, 34, 38, 42, and 46.[11]

We measure childhood socio-emotional skills at age-10 using the 'Education Questionnaire' completed by teachers, which contains 58 items from the Rutter, Conners, and Swansea scales.[12] This questionnaire includes questions on, for example, the extent to which the child "is prone to daydreaming". After excluding 16 items, primarily motor-skill related, we use the remaining 42 behavioural items alongside eight age-10 cognitive test scores spanning reading, writing, matrix and verbal reasoning to construct our measures of skills. In supplementary analyses, we complement these teacher-reported measures with child-reported measures of locus of control (CARALOC; Gammage, 1975) and self-esteem (LAWSEQ; Lawrence, 1981). Family background, measured by variables collected at age 10 or earlier, is captured by: number of siblings, birth order, father present at birth, teenage mother, family income, and parental employment and education.

At age 16, we measure participation in competitive sports from the 'Health-related Behaviour Questionnaire', where we sum the total number of sports at which teens reported representing either their school or a club over the previous year. We use five items from the 'Friends and Outside World (Document H)' and seven items from 'Attitudinal Scale - At Leisure (Document C)' to extract information on socialization. To capture mental health, we use the total score from the 22 questions of the Malaise Inventory, developed from the 'Cornell Medical Index Health Questionnaire' (Brodman et al., 1951). Finally, we obtain a measure of career interests from the 'JIIG-CAL Occupational Interests Questionnaire'.[13] In the main analysis we focus on career interests in *business* occupations as capturing an interest in highly paid jobs. Additional aspects of career orientation are examined in supplementary results.

At later ages, we use information on earnings, hours of work, and three-digit occupational codes collected every four years from ages 26 to 46 (1996 to 2016), providing a maximum of 6 observations per respondent over early to mid career. We measure completed education as that recorded at age 26.

We provide further details on the construction of our measures and the scales from which they are derived in Appendix A. Table A.1 shows summary statistics for all the variables used in the analysis.

[11] For more information on the BCS study see https://cls.ucl.ac.uk/cls-studies/1970-british-cohort-study.

[12] These comprise the 'Child's Social Behaviour' (Section B) and 'Child's Development Behaviour' (Section C) sections. Teachers rated each item on a continuous scale, coded as integers from 1 to 47.

[13] Our measures are derived using the Closs algorithm (Closs, 1978).

3.2 Sample

We start with 14,870 individuals observed at age 10, and exclude 5,525 with multiple missing items within any scale. For individuals with at most one missing item per scale, we impute the missing value using a random forest model.[14] We further remove an additional 2,393 children with no labour market information from age 26 to 46, yielding a main sample of 6,952 individuals (23,451 individual-year observations). Adding age-16 socialization reduces the sample to 3,377; further requiring occupational interests reduces it to 1,847. Throughout the analysis, we use unified samples within each specification to ensure comparability. Sample construction details are provided in Table A.2.

4 Methodology

4.1 Factorization

Following the recent economic literature on child development, we interpret the items in our dataset as multiple measurements of latent unobserved factors, and use the machinery of factor analysis to recover them. This approach has several attractive features: it addresses the conceptual question of assessing by how many dimensions, and in what way, individuals differ; it reduces hundreds of potential explanatory items to a small set of distinct factors; and it provides a principled way to handle measurement error. Our analysis follows the approach of Heckman et al. (2013) and Bolt et al. (2021); we provide an overview here and refer the reader to Appendix A.5 for technical details.

We model each standardized item as a noisy measure of an underlying latent factor:

$$Z_{\omega,i,j} = \lambda_{\omega,j}\omega_i + \epsilon_{\omega,i,j} \tag{1}$$

where $Z_{\omega,i,j}$ is the response to item j for individual i; $\lambda_{\omega,j}$ is the factor loading; ω_i is the latent factor (normalized to mean zero, unit variance); and $\epsilon_{\omega,i,j}$ is mean-zero measurement error, assumed uncorrelated across items and with the factor.

The number of factors and the assignment of items to factors are determined by exploratory factor analysis (EFA), grouping items into three broad domains: cognitive and socio-emotional skills; family background; and teen (age-16) socialization. Within each domain, we follow an iterative procedure by dropping items with loadings below 0.4 or with high cross-loadings, and retaining factors with eigenvalues above one (the Kaiser criterion). This terminates when a stable solution is reached. This yields a system where each factor is aligned with a specific set of items,

[14] We provide more detailed information on the imputation and present validation statistics in Appendix A, Table A.3. Our main results, discussed in Section 6, are robust to excluding individuals with missing items, and are available upon request.

making the empirical content of each factor highly transparent. We then estimate the final factor system using confirmatory factor analysis (CFA), setting all cross-loadings to zero, and compute individual-level Bartlett factor scores.

4.2 Long-term Outcomes

Our specification for long-term outcomes takes the following linear form:

$$y_{i,t} = \Gamma\Omega_{i,\langle a\rangle} + \beta^l X^l_{i,t} + \beta^f X^f_i + v_{i,t} \tag{2}$$

where y is the outcome of interest for individual i measured at age t, Ω captures skills measured at age $\langle a\rangle$ (10 and/or 16), X^f captures family background confounders, and $X^l_{i,t}$ captures contemporaneous adult controls. We control flexibly for differential life-cycle profiles by including an interaction of age (wave-year) with gender.

Our main outcomes include *years of schooling*, a time invariant measure obtained from the highest qualification obtained by age 26, and (log) *earnings*, measured from age 26 to age 46. We also look at (log) wages, hours of work, and being in employment. When considering labour market outcomes, we do not control for years of schooling, since this is likely a mediator of the influence of childhood skills. We document the mediating role of education for some of our headline results by showing how age-10 skill coefficients change when years of schooling are included.[15]

Since skills $\Omega_{i,\langle a\rangle}$ are not observed directly, we replace them with the Bartlett scores. This yields a linear errors-in-variables model where the signal-to-noise ratio is directly estimable from the CFA. We correct for attenuation bias using the bias-correction factor derived by Heckman et al. (2013), specified in Appendix A.6. We cluster standard errors at the individual level, allowing for serial correlation in earnings across years that is not captured by the main model features.

5 Factorization Results

As discussed in Section 4, we use an iterative exploratory factor analysis (EFA) to assess dimensionality and search for a system of dedicated measures for each factor. Our main focus is on age-10 skills. We then perform separate analyses for family background, and age-16 socialization. Here we discuss the most important results, with more details to be found in Appendix B.

[15]We also explored the role of differential life-cycle profiles by education by adding interactions between years of schooling and age or marital status, but the results changed only marginally.

5.1 Age-10 Skills

At age 10 we start with a total of 50 items comprising eight cognitive test scores and 42 measures which are taken from the sections of the 'Education Questionnaire' devoted to child behaviour.[16] After iterating the factorization and dropping items with loadings below 0.4 or high cross-loadings, we retain five factors with adjusted eigenvalues equal to or larger than one. As validation, the scree plot and Horn's parallel analysis are shown in Figure B.1.[17]

Table 1 shows results from the first round of EFA after oblique rotation, which minimizes cross-loading of items on factors and allows underlying factors to be correlated. Drawing parallels with the modern Strengths and Difficulties Questionnaire (SDQ), we label the socio-emotional factors as relating to problems of 'attention', 'conduct', 'emotion' and 'peer', with higher scores indicating more problems and lower regulation skills. The final factor we label 'cognition'. Measures that do not survive the procedure are shown in plain (non-bold) font; overall we drop 8 of the 50 items. The bottom row shows that Cronbach's alpha is well above 0.8 for all the factors, indicating high internal consistency.

Our factorization relates closely to the scales from the SDQ (Goodman, 1997), with the correspondence shown in Table B.1. The main differences relate to our first factor (attention): several items that match the SDQ's hyperactivity scale (such as 'easily distracted') load on this factor, but other items that align with hyperactivity in the SDQ load instead on conduct in our model (such as 'excitable'). For this reason we give our first factor the label '(problems of) attention' rather than hyperactivity. Otherwise, the scales line up closely, with the SDQ conduct, emotion and peer problem dimensions matching our equivalent factors.

Our four-factor representation of socio-emotional skills differs from recent papers in economics, which, using the Rutter questionnaire (Attanasio, De Paula, & Toppeta, 2020) or the Bristol Social Adjustment Guide (Papageorge et al., 2019), find two factors, typically labelled 'externalising' and 'internalising'. The main difference is due to the fact that we start our analysis from a much broader set of items - 42 versus 11 in Attanasio et al. and 10 in Papageorge et al. - drawn from multiple psychological instruments (Rutter, Conners, and Swansea scales). These items provide richer measurement across behavioural domains, and allow us to identify attention problems as distinct from conduct problems, and peer problems as distinct from emotion problems. The four-factor representation is also supported by recent psychometric studies of the SDQ structure, which indicate that a finer factorization displays better goodness of fit than a coarser factorization (Bell et al., 2024; Bridger-Staatz et al., 2024). One of our main contributions is to find that this finer factorization matters: attention and conduct problems, typically combined under 'externalising',

[16] As discussed previously, in a prior step we exclude 16 items related to fine motor skills, such as hand-eye coordination.

[17] Parallel analysis (Horn, 1965) adjusts eigenvalues for sampling noise using a simulation-based approach, and remains one of the standard approaches for determining factor dimensionality (Lim and Jahng 2019). See Appendix B for an explanation of other technical terms used in this section.

Table 1: Exploratory Factor Analysis (EFA) of Age-10 Skills: First Iteration

Items	Attention Problems	Conduct Problems	Emotional Problems	Peer Problems	Cognition
Easily distracted	**0.793**	0.137	0.029	-0.116	-0.043
Fails to finish tasks	**0.783**	-0.033	-0.029	0.064	-0.006
Cannot complete tasks	**0.773**	-0.050	-0.058	0.072	-0.030
Fails to pay attention in class	**0.725**	0.056	-0.086	0.069	-0.095
Fails to show perseverance	**0.722**	0.023	-0.062	0.060	-0.084
Bored in class	**0.690**	0.159	-0.018	0.074	-0.022
Daydreaming	**0.671**	-0.128	0.112	0.124	0.022
Forgetful on complex task	**0.611**	-0.039	0.211	-0.010	-0.203
Cannot concentrate on task	**0.584**	0.004	0.041	0.027	-0.082
Squirmy and fidgety	0.543	0.388	0.122	-0.183	0.039
Shows lethargic behaviour	**0.493**	-0.020	0.138	0.290	0.001
Confused with diffic. tasks	0.457	-0.085	0.388	-0.031	-0.266
Displays outbursts of temper	-0.053	**0.798**	0.037	0.068	-0.030
Teases other children	0.030	**0.783**	-0.125	0.013	-0.036
Bullies other children	-0.025	**0.781**	-0.137	0.101	-0.073
Quarrels with other kids	0.041	**0.759**	-0.005	0.141	-0.079
Changes mood quickly	0.024	**0.701**	0.253	-0.011	-0.014
Interferes with others	0.289	**0.643**	-0.104	-0.022	-0.026
Complains about things	0.052	**0.623**	0.125	0.020	-0.030
Sullen or sulky	0.015	**0.623**	0.101	0.255	-0.039
Destroys belongings	0.099	**0.609**	-0.046	0.142	-0.013
Excitable and impulsive	0.124	**0.597**	0.217	-0.345	0.023
Restless or over-active behv.	0.295	**0.554**	0.215	-0.247	0.056
Easily frustrated	0.164	**0.526**	0.207	-0.068	-0.017
Hums or makes odd vocals	0.277	**0.436**	0.012	-0.043	0.048
Rhythmic tapping in class	0.289	**0.419**	0.037	-0.056	0.056
Cannot negotiate child's behv.	0.281	0.316	-0.126	0.278	0.042
Face or body twitches	0.115	0.251	0.199	0.019	0.045
Worried	0.000	0.061	**0.829**	0.005	-0.013
Behaves nervously	0.127	-0.055	**0.719**	0.073	-0.033
Anxious	-0.037	-0.024	**0.699**	0.182	-0.031
Fussy	-0.093	0.300	**0.564**	-0.050	0.043
Afraid of new situations	0.162	-0.177	**0.560**	0.103	-0.172
Obsessed with unimportant tasks	0.015	0.377	0.447	0.050	0.003
Cries for little cause	-0.048	0.333	0.416	0.087	-0.042
Unhappy or tearful	-0.012	0.355	0.371	0.351	-0.006
Child is not friendly	0.141	0.172	0.057	**0.719**	-0.035
Child is not popular with peers	0.160	0.227	0.028	**0.698**	-0.043
Introvert	-0.035	-0.316	0.297	**0.606**	-0.005
Rather solitary	0.017	0.047	0.243	**0.549**	0.100
Child is not cooperative	0.147	0.365	-0.046	**0.538**	-0.026
Child is not bold	0.016	-0.434	0.333	0.450	-0.069
BAS words	0.094	-0.025	0.008	-0.001	**0.806**
Reading	-0.113	-0.006	0.024	0.015	**0.800**
Maths	-0.087	0.012	-0.017	0.006	**0.776**
Pictorial (PLC)	0.109	-0.059	-0.001	0.018	**0.773**
BAS simil	0.089	-0.019	-0.006	0.015	**0.760**
BAS matrix	-0.053	-0.043	0.032	0.018	**0.629**
Spelling	-0.230	0.065	-0.010	0.011	**0.545**
BAS digits	-0.074	0.029	-0.022	0.009	**0.420**
Cronbach's alpha	0.925	0.931	0.847	0.823	0.856

Notes: Data from BCS70. The table reports factor loadings obtained from an exploratory factor analysis (EFA) of the main sample (6952 observations) and oblique quartimin rotation. Items in bold are subsequently retained after several iterations and used in our dedicated measurement system. The items not in bold are dropped. Cronbach's alpha coefficients for the internal consistency of the set of retained items within each factor are also reported.

have opposite relationships with economic outcomes, and these distinctions remain hidden in the coarser two-factor models prevalent in previous studies.[18]

More broadly, it is instructive to relate our findings to the large literature examining the role of personality traits in labour market outcomes. The dominant framework has been the 'Big Five' or 'OCEAN' model (Almlund et al., 2011; McCrae & Costa, 1999); the economics literature has also used bespoke frameworks, including, influentially, those provided by military assessments in Nordic countries. In Appendix D we relate our behaviour measures to personality traits using the available evidence. We argue that our finding of a positive return to conduct is consistent with these alternative frameworks, particularly because of its correlation with (dis-)agreeableness and with extraversion. However, the measures we use here have fundamental differences compared to those from personality psychology. First, they are designed for children and have their origins in clinical psychology as screens for psychopathologies; as such they are more directly relevant to school policies and other interventions. Second, the behavioural measures are provided by external observers (teachers and mothers) rather than by self-reports. As we discuss in Section 6, the identity of the respondent matters for the predictive quality of the measures.

Further exercises support our preferred factor structure. A gender-specific EFA yields very similar loadings across men and women, with Cronbach's alpha remaining above 0.8 for all factors (Table B.3). A multitrait-multimethod (MTMM) analysis following Goodman et al. (2010), which exploits the fact that a subset of items are reported by both teachers and mothers, supports convergent and discriminant validity of the four socio-emotional factors. Here, the weakest discrimination arises between attention and conduct, a feature also documented for the SDQ in Goodman et al. (Table B.4). A separate factorization of mother reports alone recovers the attention, conduct, and emotional factors but not peer problems, consistent with the MTMM results which show lower internal validity for the mother-reported constructs (Table B.5). Finally, including 16 items from the CARALOC locus-of-control scale and 12 items from the LAWSEQ self-esteem scale does not convincingly yield an additional factor: locus-of-control items load weakly and are not retained, while self-esteem items generate a factor with Cronbach's alpha only just above 0.7 (Table B.6). We present additional results including self-esteem as a robustness check in Appendix C.

5.2 Family Socio-Economic Status and Age-16 Socialization

We perform a separate factorization for questions about family socioeconomic status such as parental education, parental employment and family income. As shown in Table B.7, all these items have loadings greater than 0.6 and are captured by a single factor with an eigenvalue of 2.1.

[18]To further investigate the relationship of our factors with the two-factor structure employed in previous work, we perform a factorization using only the Rutter items and find, consistently with these studies, that only two factors are retained (Table B.2).

To capture socialization at age 16, we use information from both the 'Friends and Outside the World' questionnaire and 'Attitudinal Scales - At Leisure'. The former provides information on the social behaviour of the teenager during school term such as 'spend time at friends' homes' and 'go out with friends do nothing special'. The latter questionnaire shows analogous measures of social behaviour during leisure time. We select 12 items including the number of friends in and outside school, and drop 4 items that load poorly. The first-round EFA for the age-16 characteristics is also shown in Table B.7.

5.3 Confirmatory Factor Analysis and Raw Correlations

We perform a Confirmatory Factor Analysis (CFA) using a dedicated measurement system as described in (1). As the sample decreases dramatically when using information at age 16, we run CFA for the full sample (excluding the age-16 items used to capture teen socialization) and the smaller sample (including those items) separately. Final loadings are given in Tables B.8 and B.9, which show similar weights across samples and good model fit on the Tucker-Lewis and Comparative Fit indices. From the main CFA we extract six Bartlett-corrected factor scores (one cognitive, four behavioural and one family SES) as given by equation (3) in Appendix A; from the CFA using additional age-16 data we extract seven scores, now additionally including teen social behaviour.

Table B.10 shows the correlation matrix of raw factor scores and additional variables of interest, computed using the narrower sample including measures at age 16. Years of schooling correlates positively and strongly with cognition and family socio-economic status, and negatively with the four socio-emotional problem behaviours. Cognition is moderately negatively correlated with attention problems and weakly with the other three behaviours. The four socio-emotional problem behaviours are positively related, with the strongest correlation between attention and conduct. Teen socialization is weakly related to all other variables, with its strongest correlation being, as expected, with age-10 peer problems. Table B.10 also displays correlations with scores from the alternative externalising/internalising factorization: again as expected, externalising correlates most strongly with conduct, then attention; internalising correlates most strongly with emotional problems, then peer problems.

6 Results on Long-Run Outcomes

Having established a factorization based on four socio-emotional skills, we now explore the relationship between these skills and later outcomes. Given that we have multiple childhood skills and multiple adult outcomes, we streamline the discussion by synthesizing the results and only discussing key relationships.

6.1 Basic Relationship With Key Outcomes

We first examine the role of early socio-emotional skills in determining total years of schooling, while controlling for key confounding characteristics and additional background variables, as per the regression specification in Section 4. Results are displayed on the left hand side of Table 2. The first column shows that, when we don't control for cognitive scores, the negative association between schooling and attention problems is extremely strong. Moving from two standard deviations below the mean to two standard deviations above is associated with over 2.5 fewer schooling years. As we discuss in Appendix D, in the language of personality psychology, attention problems are aligned with low conscientiousness, which is often found to be positively related to educational attainment (Almlund et al., 2011). The first column also suggests a positive effect of conduct on schooling, but little role for emotional or peer problems.

Controlling for cognitive scores (column 2) changes the interpretation substantially. The association between schooling and attention is reduced by around two thirds. Meanwhile, the apparent positive effect of conduct problems switches sign and is no longer significant. On the other hand, the positive coefficient on emotional problems increases and is now significant at the 10% level.

The right hand side of Table 2 focuses on the relationship between childhood behaviours measured at age 10 and earnings, measured between ages 26 and 46. The third column includes measures of socio-emotional skills and controls for the life-cycle profile of earnings and family socio-economic status, but does not take into account cognition.[19] Similarly to Prevoo and ter Weel (2015), the data show a strong negative relationship between problems of attention and earnings. Moving from two standard deviations above the mean for this behaviour to two standard deviations below is associated with an increase in earnings of 37 log points, or around 45%.

The next row shows the most striking result from this table. Bad conduct is associated with strong *positive* effects on earnings. We provide a preliminary explanation for this result by noting that problems with conduct most saliently include aggression, which in later life has consistently been found to relate to positive returns in the labour market.[20] Moving on to the remaining behavioural skills, and consistent with earlier studies, emotional problems are negatively associated with earnings, conditional on the other behavioural factors. Finally, in this initial analysis, peer problems seem to have no noticeable relationship with earnings.

Column 4 presents our benchmark specification, which additionally controls for cognitive ability. Cognition itself has a strong association with earnings: a one-standard deviation increase is associated with a little over 9% higher earnings. More relevantly, including it in the controls

[19]We include family SES as a single factor, as this allows an easy comparison of its effect size with those for the behaviours of interest. When we replace the single factor with a full set of disaggregated background family controls, results are near identical.

[20]See, for example, Almlund et al. (2011). In more detail, and as we explore in Appendix D, these behaviours are associated with reduced agreeableness in adulthood; as agreeableness correlates negatively with labour market earnings across a range of measures, this is consistent with the positive returns documented here.

Table 2: Determinants of Schooling and Earnings

	Schooling		Earnings		
	[1]	[2]	[3]	[4]	[5]
Attention	-0.648^{***}	-0.221^{***}	-0.092^{***}	-0.037^{***}	-0.027^{***}
	[0.038]	[0.045]	[0.008]	[0.009]	[0.009]
Conduct	0.094^{***}	-0.037	0.052^{***}	0.036^{***}	0.037^{***}
	[0.036]	[0.036]	[0.007]	[0.007]	[0.006]
Emotion	0.025	0.058^{*}	-0.032^{***}	-0.028^{***}	-0.029^{***}
	[0.035]	[0.033]	[0.007]	[0.007]	[0.006]
Peer	0.040	0.025	-0.010	-0.011	-0.014^{**}
	[0.035]	[0.032]	[0.007]	[0.007]	[0.007]
Cognition		0.726^{***}		0.093^{***}	0.060^{***}
		[0.036]		[0.007]	[0.007]
Family SES	0.856^{***}	0.610^{***}	0.107^{***}	0.075^{***}	0.046^{***}
	[0.032]	[0.032]	[0.007]	[0.007]	[0.007]
Yrs School					0.049^{***}
					[0.003]
Background controls	X	X	X	X	X
Mean of Dep. Var.	12.26	12.26	7.17	7.17	7.17
N individuals	6,952	6,952	6,952	6,952	6,952
N individual-years			23,451	23,451	23,451

Notes: Data from BCS70. Each column reports measurement-error-corrected estimates from a regression of log monthly earnings and years of schooling on standardized socio-emotional skills, cognition, and family socio-economic status (SES) obtained from our dedicated measurement system (see equation (1)). All specifications control for: gender, number of siblings, dummies for first child, no dad at birth, teenage mother, and region fixed effects. Specifications in the earnings model also include controls for gender-by-year and year fixed effects. Standard errors in square brackets are estimated from 250 bootstrap replications, clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

modifies the observed relationship between earnings and behavioural skills. Most notably, and similarly to the schooling regression, the coefficient on attention is reduced by around 60%. Likewise, the coefficients on conduct and emotional problems are now smaller in magnitude, although they remain strongly significantly different from zero.

In the final column we examine the role of years of schooling as a mediator. Given the strong positive correlations between education and cognitive scores and background family SES (including parental education), we expect to see changes in those coefficients, and indeed we see that the relationship of these two characteristics with earnings reduces significantly as part of their raw effect is now mediated by schooling. In terms of our variables of focus, the coefficient on attention problems is also brought down by over a quarter. However, controlling for schooling if anything *strengthens* the estimated relationships with the other socio-emotional skills. In particular, the coefficient on peer problems becomes significant at the 5% level, suggesting that schooling, if anything compensates for problems in this area.

Overall, the message from the right hand side of Table 2 is clear and strong: behaviours that are deemed negative in terms of attentiveness and emotional stability have negative associations with long-term labour market prospects, while aggressive conduct has apparent positive effects.

We further explore the relationship with earnings by breaking it down into wages and hours, as well as examining labour supply at the extensive margin and looking at the probability of being in employment. Table 3 shows that the associations of childhood skills with wages are the same sign as in the earnings regressions, shown previously. It also shows a more pronounced negative role for peer problems in terms of hourly wage rates. Comparing determinants of wages and hours in columns 1 and 2, perhaps the most noticeable finding is that the negative effect of emotional problems on earnings captures both reduced wages *and* reduced hours of work. Similarly, the positive effect of conduct on earnings partly reflects increased hours of work, though in this case mainly operates through a wage premium.

Column 3 of Table 3 displays results at the extensive margin (employment) and shows subtly different effects compared to the intensive margin (hours).[21] Attention problems lead to noticeably *lower* engagement with the labour market. However, there seems to be no overall association for emotional and conduct difficulties, though peer problems are weakly predictive of slightly lower engagement with the labour market overall. As expected, labour market engagement is positively associated with family background and cognition.

The combined results across outcomes shown so far are worth discussing in relation to Papageorge et al. (2019), who find that some child behavioural traits are associated with outcomes of opposing signs in the labour market versus at school. Papageorge et al. draw from this the policy implication that schooling may not be rewarding skills appropriately, putting children with valuable labour market skills at risk of being cast aside when young. Specifically they find this pattern for externalising behaviours, one of the two dimensions of socio-emotional skills in their representation, alongside internalising behaviours. As discussed in Section 5, externalising behaviours combine attention and conduct problems, while internalising behaviours reflect emotional and peer problems.

To explore this aspect further, in Table 4 we show how our results - based on a four-factor representation of skills - compare to those obtained from a two-factor model - which is more commonly

[21]It is worth noting that in analysing the probability of being in employment in column 3 of Table 3, we consider both employees and the self-employed. The rest of our analysis looks at employees only. When we repeat the analysis excluding the self-employed, results are similar. Also, to emphasize, our main results for earnings shown in Table 2 condition on positive labour supply only.

Table 3: Determinants of Wages, Working Hours and Employment

	Wages [1]	**Hours** [2]	**Employment** [3]
Attention	-0.037^{***}	-0.001	-0.015^{***}
	[0.007]	[0.005]	[0.005]
Conduct	0.025^{***}	0.010^{***}	-0.003
	[0.005]	[0.004]	[0.004]
Emotion	-0.011^{**}	-0.017^{***}	-0.003
	[0.005]	[0.003]	[0.004]
Peer	-0.016^{***}	0.005	-0.006^{*}
	[0.006]	[0.004]	[0.003]
Cognition	0.077^{***}	0.016^{***}	0.024^{***}
	[0.006]	[0.004]	[0.004]
Family SES	0.068^{***}	0.007^{*}	0.013^{***}
	[0.005]	[0.004]	[0.004]
Background controls	X	X	X
Mean of Dep. Var.	2.16	5.02	0.85
N individuals	6,952	6,952	8,162
N individual-years	23,451	23,451	34,613

Notes: Data from BCS70. Each column reports measurement-error-corrected estimates from a regression of log hourly wages (column 1), log monthly working hours (column 2), and whether an individual is employed (employee or self-employed; column 3) on standardised socio-emotional skills, cognition, and family socio-economic status (SES), constructed using our dedicated measurement system (see equation (1)). The employment outcome (column 3) is estimated via a linear probability model (LPM) using all survey participants with valid observations across the 1996–2016 waves. All specifications include controls for gender, gender-by-year, number of siblings, dummies for first child, no father at birth, teenage mother, and year and region fixed effects. Standard errors in square brackets are estimated from 250 bootstrap replications, clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

used in other studies (see also Attanasio, Blundell, et al., 2020). The two-factor representation replicates Papageorge et al.'s findings: externalising relates negatively to schooling but positively to earnings and hours of work. However, our four-factor model reveals these contrasting results stem from different behaviours. The negative schooling effect is driven mostly by attention problems, while the positive labour market association is driven mostly by the positive relationship with conduct problems. This distinction resolves the apparent contradiction and shows that the reasons why some children struggle in school are distinct from the behavioural skills that are beneficial in the labour market. This finding suggests that policies aimed at improving educational attainment should mainly target attention problems, which are the main driver of poorer schooling outcomes. At the same time children who display disruptive classroom behaviour should not be penalised, as those behaviours might reflect skills - such as assertiveness or competitiveness - that carry a labour market premium.

It is worth noting that we do see this feature of opposing outcomes for a different set of behaviours: emotional problems. Children who score higher on emotional difficulties tend to do worse in the labour market, but also to stay in school longer - significant at the 10% level. Here the policy implications depend on whether the returns to education reflect genuine skill accumulation or mainly signalling. If the former, this result is encouraging as it shows that the schooling system is helping anxious and vulnerable children accumulate human capital that partially offsets their later disadvantage at work. If the latter, an extra year of schooling generates credentials that overstate these individuals' productivity in the workplace, and this is inefficient. The right policy message therefore depends on an empirical question: whether the returns to education reflect signalling vs. human capital, and this is still a matter of debate (see for example Ehrmantraut et al., 2020).

6.2 Additional Results on Labour Market Outcomes

In Appendix C we provide evidence on the extent to which the associations between childhood skills and earnings operate through sorting into different occupations. Figure C.1 shows that children with conduct behaviours are more likely to sort into management and associate professional jobs. Strikingly, we also see that those with attention problems are less likely to work in associate professional or administrative jobs but are more likely to work in trades and sales, possibly as they prefer not to be in an office environment. Focusing on conduct problems, Figure C.2 shows in more detail that children who display more of these behaviours sort into jobs that require higher stress tolerance, leadership and a willingness to compete, according to O*NET descriptors.[22] When we control for 3-digit occupation and schooling in earnings regressions, we see that these variables mediate some of the effects of childhood skills, though significant associations remain for conduct and emotional problems in particular (Table C.1).

Appendix Tables C.2 and C.3 provide evidence on gender differences. Overall, most effects for our key skills of focus are not significantly different across genders, with a few noteworthy exceptions. There is a stronger negative effect of attention problems on schooling for boys than girls. In terms of earnings we see a stronger negative effect of peer problems for boys, driven by differentially lower labour supply at the intensive margin, and indicating more part-time rather than full-time work. When looking at the extensive margin (employment), we see that the effect of attention on labour market attachment is driven mainly by girls. Conduct problems are associated with increased employment for girls, but reduced employment for boys. The extent to which

[22]We thank an anonymous referee for suggesting this analysis.

Table 4: Main Regressions Comparing 4-Factor with 2-Factor Models

	Schooling		Earnings		Wages		Working Hours	
	4-factor	**2-factor**	**4-factor**	**2-factor**	**4-factor**	**2-factor**	**4-factor**	**2-factor**
Externalising		-0.143^{***}		0.020^{***}		0.008		0.012^{***}
		[0.034]		[0.007]		[0.005]		[0.003]
Attention	-0.221^{***}		-0.037^{***}		-0.037^{***}		-0.001	
	[0.045]		[0.009]		[0.007]		[0.005]	
Conduct	-0.037		0.036^{***}		0.025^{***}		0.010^{***}	
	[0.036]		[0.007]		[0.005]		[0.004]	
Internalising		0.041		-0.046^{***}		-0.032^{***}		-0.014^{***}
		[0.031]		[0.008]		[0.006]		[0.004]
Emotion	0.058^{*}		-0.028^{***}		-0.011^{**}		-0.017^{***}	
	[0.033]		[0.007]		[0.005]		[0.003]	
Peer	0.025		-0.011		-0.016^{***}		0.005	
	[0.032]		[0.007]		[0.006]		[0.004]	
Cogn. & Fam SES	X	X	X	X	X	X	X	X
Backg. Controls	X	X	X	X	X	X	X	X
Mean of Dep. Var.	12.26	12.26	7.17	7.17	2.16	2.16	5.02	5.02
N individuals	6,952	6,952	6,952	6,952	6,952	6,952	6,952	6,952
N individual-years			23,451	23,451	23,451	23,451	23,451	23,451

Notes: Data from BCS70. Each column reports measurement-error-corrected regression estimates comparing a four-factor model of socio-emotional skills with a two-factor model (internalising and externalising); see Table B.2. Socio-emotional skills, cognition, and family socio-economic status (SES) are constructed using our dedicated measurement system (see equation 1). All specifications include controls for gender, gender-by-year, number of siblings, dummies for first child, no father at birth, teenage mother, and year and region fixed effects. Standard errors in square brackets are estimated from 250 bootstrap replications, clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

these gender differences are driven by differences in family composition and fertility choices is not something we explore further here.

Appendix C also presents robustness of our key results to a range of additional analyses. (i) We examine potential selection bias from differential survey attrition based on socio-emotional skills. While individuals with emotional problems are more likely to remain in the survey and those with conduct problems are more likely to drop out, inverse probability weighting to correct for this differential attrition has minimal impact on our estimates (Tables C.4 and C.5). (ii) Including self-esteem measures from the LAWSEQ scale similarly leaves our main results unchanged, with self-esteem showing only modest independent effects on labour market outcomes (Table C.6). (iii) We find little evidence of non-linearities in the relationships between childhood skills and adult outcomes, with effects remaining monotone across the main support of our explanatory variables

(Figure C.3). (iv) We examine a range of interactions across skills. We find some evidence of interactions between our socio-emotional skills (Table C.7). Most notably, we see evidence of complementarity between conduct and attention, such that conduct has a higher return at *lower* levels of attention problems. We also find modest complementarity between cognitive skills and the capacity for attention (Table C.8). However, there is no evidence of significant interactions with the other age-10 socio-emotional measures. In all these specifications, the negative relationship between earnings and problems of attention and the positive relationship with problems of conduct remain robust.

Finally, we show the strength of our teacher-based factor measures by comparing them with analogous factors derived from mother reports. Table C.9 shows that while mother-assessed skills generally align with teacher assessments in terms of the direction of the effects, they are consistently weaker predictors of both schooling and earnings outcomes. The one exception is emotional problems and schooling, where mother-reported items show stronger associations, likely reflecting greater visibility of these behaviours at home. When both teacher and mother assessments are included simultaneously, teacher reports dominate for most relevant variables, demonstrating their superior predictive validity for adult long-term outcomes. This reinforces our focus on teacher-assessed measures as the primary basis for our four-factor model.

6.3 Evidence on Causal Interpretation

Our evidence so far relies on associations between childhood skills and later outcomes and cannot be interpreted as causal. The fact that our skills are measured at age 10, well before labour market entry, eliminates the possibility of reverse causality but does not rule out other endogeneity concerns. Childhood skills and later economic success could both be simultaneously influenced by unobserved family characteristics, such as parenting quality, family values, or parental networks. Similarly, neighbourhood effects or school quality might jointly determine both observed skills and subsequent outcomes.

We address these concerns through two complementary approaches. First, in Table C.10 we show that our estimates are robust to the progressive inclusion of controls for neighbourhood characteristics, housing quality, parenting style, early-life health, and school characteristics. If anything, these additional controls slightly *strengthen* the estimated effects. Second, and perhaps more persuasively, we exploit sibling variation in additional data, which allows us to control for shared family-level confounders.

Since the BCS70 sample design yields very few sibling pairs, for this second exercise we turn to the NLSY79 Child and Young Adult survey (NLSY79 CYA), which tracks the children of female respondents in the original NLSY79 cohort. This dataset is well-suited for our analysis because it contains behavioural assessments from the Behavior Problems Index (BPI) that can be mapped to our BCS70 measures, and its family-based sampling frame provides approximately 3,000 sibling pairs. Our core analytical sample, for which BPI and labour market outcomes are observed, comprises 5,150 individuals born between 1971 and 1994 from roughly 1,900 sibling groups, yielding 26,238 total earnings observations spanning 2000 to 2020. While the NLSY79 CYA sample differs from our BCS70 cohort in some ways (participants are on average younger and more likely to have been born to teenage mothers with larger family sizes) we can control for these demographic differences in our regression framework.[23]

Our first task is to estimate factors corresponding to our skills. Given the more limited set of items in the BPI inventory, an exploratory factor analysis shows less variation than for the items in the BCS70 and picks out only two factors. To ensure comparability with our benchmark analysis, we therefore impose a four-factor structure that mirrors our BCS70 approach.

Specifically, we use all the BPI items available to extract the first four principal components, and then rotate these to align with our conceptual framework. We rotate components using exclusion restrictions based on theoretically motivated item-factor mappings. Specifically, we constrain four key items to load exclusively on their corresponding factors: 'Has difficulty concentrating/paying attention' loads only on the first factor (Attention), 'Argues too much' loads only on the second factor (Conduct), 'Worries too much' loads only on the third factor (Emotion), and 'Has trouble getting along with other children' loads only on the fourth factor (Peer). These exclusion restrictions serve as identification anchors, ensuring that each factor captures its intended behavioural dimension. For all other BPI items, we allow cross-loadings across factors. Table E.2 shows the resulting factor loadings, while Table E.3 shows the resulting correlation structure between factors. This latter table also shows the corresponding correlation structure for the BCS70 and indicates strong similarity in correlations between the two sets of factors, validating our approach.

We then use these factors to investigate the relationship with later outcomes. The results for schooling and for earnings are shown in Table 5. Starting with schooling, the first column of panel A reproduces the BCS70 results from Table 2, while the second column shows the corresponding estimates using the NLSY79 CYA sample. To improve comparability with our benchmark results, this specification uses a slightly wider sample than our core NLSY79 CYA sample, including an extra 878 children with no information on siblings. With the exception of peer problems, the coefficient signs line up fairly well with those from the BCS70. Effect sizes are stronger for conduct, which now has a negative significant association with schooling, and for emotional problems. The differences are less quantitatively important for attention and cognition.

The third column of Panel A restricts the analysis to the sibling sample, and shows almost identical results. Column 4 allows for sibling fixed effects. Here, estimates for attention and peer problems become slightly smaller than in column 3, although not significantly so. The main change is a large reduction in the coefficient for conduct problems, which is no longer significantly asso-

[23] A detailed discussion of the NLSY79 CYA, including its strengths and limitations, is provided in Appendix E, with comparative summary statistics presented in Table E.1.

ciated with schooling, and in the coefficient for emotional problems, which is now only weakly significant. This suggests that much of the population-wide association between conduct and emotion problems and educational attainment may be attributable to family-level confounding factors, plausibly reflecting that families with higher levels of these behaviours place less emphasis on education overall. Similarly, for cognitive skills, we see that the within-family estimate is much smaller in magnitude, reflecting the presence of relevant family-level confounders.

A critical methodological issue is that measurement error complicates the interpretation of sibling fixed effects estimates, amplifying attenuation bias. (See Bound & Solon, 1999; Fletcher, 2013; Holmlund et al., 2011; Sjölander et al., 2022). Addressing this issue is beyond the scope of the current study. Instead we note that the literature suggests that, due to this consideration, the reduction in size of estimates between columns 3 and 4 potentially represents an upper bound on the reduction of true effect size owing to confounding by family-level factors.[24]

Panel B shows equivalent results for earnings. Two important findings emerge. First, for attention, cognition and notably for conduct the effect of skills on earnings matches quite closely that in the BCS70. Even for emotion and peer problems, the point estimates are of the same sign, with the difference in magnitude and significance perhaps attributable to the fact that in the NLSY79 CYA these measures are more difficult to separate given the more limited set of items available.

Second, and in contrast to the schooling results, we find little to no evidence of confounding by family-level factors. The sibling fixed effects estimates in column 4 are statistically indistinguishable from the pooled (not-fixed-effects) estimates in column 3, with the formal p-values for the differences always a long way above conventional significance thresholds. This pattern suggests that unobserved family characteristics play a much smaller role in determining labour market outcomes than educational attainment, plausibly reflecting that families have more direct influence over their children's schooling decisions and educational environments than over adult labour market performance.

A consideration when comparing results across datasets concerns the source of the behavioural reports. While our BCS70 factors are constructed from teacher assessments, the NLSY79 CYA factors are derived from mother reports of the BPI. As shown in Table C.9, discussed previously, within the BCS70 the conduct factor derived from mother reports is a weaker predictor of earnings than its teacher-reported counterpart, and similar differences apply to other factors. This measurement difference may account for some of the cross-dataset discrepancies. Viewed in this light, the replication of a positive and significant conduct–earnings association in the NLSY79 CYA is if

[24]Additionally, as discussed in Currie and Almond (2011) the estimates in the fourth column can be interpreted most cleanly under the assumption of no confounding *within*-family variation. For example, a plausible explanation of the results for cognitive skills is that parents put compensating investments into their less intelligent children, boosting their educational attainment. In this case the estimates here identify the effect of variation in skills in the context of this type of parental investment strategy and so do not, for example, identify the effect of raising skills uniformly across the family or across society.

Table 5: The Effect of Cognitive and Socio-Emotional Skills: BCS70 vs NLSY79 CYA

	BCS70	**NLSY79 CYA**	**NLSY79 CYA Sib. Sample**	**NLSY79 CYA Sib. Sample**	**p-val**
	[1]	[2]	[3]	[4]	[3] - [4]
Panel A: Schooling					
Attention	-0.221^{***}	-0.150^{***}	-0.141^{***}	-0.124^{***}	0.563
	[0.045]	[0.029]	[0.034]	[0.039]	
Conduct	-0.037	-0.078^{**}	-0.086^{**}	-0.028	0.095
	[0.036]	[0.033]	[0.037]	[0.043]	
Emotion	0.058^{*}	0.104^{***}	0.101^{***}	0.055^{*}	0.085
	[0.033]	[0.025]	[0.029]	[0.032]	
Peer	0.025	-0.158^{***}	-0.139^{***}	-0.124^{***}	0.700
	[0.032]	[0.033]	[0.036]	[0.043]	
Cognition	0.726^{***}	0.571^{***}	0.593^{***}	0.393^{***}	0.000
	[0.036]	[0.027]	[0.033]	[0.038]	
Mean of Dep. Var.	12.26	13.28	13.09	13.09	
N individuals	6,952	6,031	5,152	5,152	
Panel B: Earnings					
Attention	-0.037^{***}	-0.054^{***}	-0.053^{***}	-0.045^{***}	0.450
	[0.009]	[0.011]	[0.012]	[0.013]	
Conduct	0.036^{***}	0.025^{**}	0.024^{*}	0.028^{**}	0.750
	[0.007]	[0.012]	[0.013]	[0.014]	
Emotion	-0.028^{***}	-0.005	-0.002	-0.007	0.576
	[0.007]	[0.009]	[0.010]	[0.010]	
Peer	-0.011	-0.047^{***}	-0.048^{***}	-0.043^{***}	0.705
	[0.007]	[0.012]	[0.013]	[0.014]	
Cognition	0.093^{***}	0.133^{***}	0.139^{***}	0.143^{***}	0.718
	[0.007]	[0.010]	[0.011]	[0.014]	
Mean of Dep. Var.	7.17	6.99	6.99	6.99	
N individuals	6,952	6,028	5,150	5,150	
N individual-years	23,451	30,583	26,238	26,238	
Siblings FE				X	
Backg. Controls	X	X	X	X	
Family SES	X	X	X	X	

Notes: Data from BCS70 and NLSY79 CYA. The table reports regression estimates of years of schooling (Panel A) and log monthly earnings (Panel B) on standardised cognitive and socio-emotional scores. In BCS70, socio-emotional and cognitive skills are constructed using our dedicated measurement system (see equation 1). In NLSY79 CYA, socio-emotional skills are obtained from a PCA on 32 mother-reported Behavior Problems Index (BPI) items collected in childhood (see Appendix E for details). We anchor these components using exclusion restrictions of representative items: *attention* ("has difficulty concentrating/paying attention"), *conduct* ("argues too much"), *emotion* ("worries too much"), and *peer problems* ("has trouble getting along with other children"), while allowing the remaining 28 items to load freely. Columns [3] restrict the sample to biological siblings and column [4] controls for sibling fixed effects. Reported p-values test equality of coefficients between columns [3] and [4]. All specifications include controls for family socio-economic status, gender, ethnicity (NLSY79 CYA), birth order, teenage mother, and region fixed effects. Earnings specifications additionally control for gender-by-age, age (NLSY79 CYA), and year fixed effects. Standard errors in square brackets are clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

anything more striking, since it emerges despite a reporter-based attenuation that works against finding such an effect.

Taken together, and bearing in mind the measurement caveats just discussed, these results are consistent with a causal interpretation of the relationship between childhood socio-emotional skills and adult labour market outcomes. The sibling fixed effects analysis reveals that while cognitive skills' effects on schooling appear somewhat attributable to family-level factors, the returns to socio-emotional skills in the labour market remain robust. This pattern aligns with Fletcher (2013)'s finding that extraversion maintains its positive earnings effect even after controlling for family fixed effects, even if he finds that effects from conscientiousness disappear.

6.4 Pathways Through Age-16 Behaviours

We have seen that age-10 skills are strongly associated with labour market outcomes, and that they operate in different ways. Most strikingly, conduct problems (such as aggression) are found to be positively related to earnings, and this effect can be partially explained both by higher wages and increased labour supply at the intensive margin. Attention problems have strong negative effects, partly because they affect schooling outcomes. Emotionally problematic behaviours, such as anxiety, lead to reduced labour supply. At the same time problems with peers are also later associated with lower wage rates. In this section, we explore behaviours measured at age 16, examining both how they relate to our age-10 measures and how they connect to the broader social skills literature (Deming, 2017; Weinberger, 2014).

We focus on four different domains that are a priori likely to be important for labour market outcomes and about which the BCS70 sweep at age 16 contains useful data: *competitive sport participation*, as evidenced by representing the school or a club in a competitive sport event; *teen socialization*, obtained using a separate factorization on items capturing the breadth of friendship networks (described in section 5.2); *career interests*, measured through the JIIG-CAL Occupational Interests Questionnaire; and *well-being and emotional health*, measured through the Malaise Inventory. As discussed in Section 3, the response rate to the survey was lower at age 16 so here we necessarily work with smaller samples.[25]

Table 6 shows relationships between the age-10 skills and the age-16 behaviours of interest. We first note that, as expected, both peer and emotional problems are negatively associated with participation in competitive sports and teen socialization.[26] However, looking at other skills, some

[25] Additionally, because the samples for these age-16 behaviours do not coincide perfectly, we examine each feature separately rather than all jointly. Although not shown here, when we do form samples including more than one of the age-16 behaviours, results are little changed. Results available upon request.

[26] As we discuss in detail in Appendix D.3, our teen socialization and sports competition measures capture similar concepts to the social skills measures used in influential work by, for example Deming (2017) and Weinberger (2014). Teen socialization aligns with Deming's sociability measures, while sports competition overlaps with both authors' participation-based measures, though ours focuses more on competitive aspects rather than leadership roles.

striking differences emerge. Conduct problems are strongly and *positively* associated with participation in competitive sports, while they have a much weaker relationship with teen socialization. At the same time cognition predicts socialization but not competitive sport engagement. This suggests that competitive sport participation captures a specific dimension of our conduct problem factor - likely related to preference for competition - rather than general social tendencies.

Table 6: Determinants of Teen Characteristics

	Sport Competition	**Teen Socialization**	**Attitudes to Business**	**Mental Health**
	[1]	[2]	[3]	[4]
Attention	-0.051	0.025	-0.147^{***}	0.012
	[0.036]	[0.035]	[0.042]	[0.039]
Conduct	0.125^{***}	0.060^{**}	0.014	-0.039
	[0.025]	[0.027]	[0.034]	[0.029]
Emotion	-0.078^{***}	-0.057^{**}	0.021	-0.071^{***}
	[0.023]	[0.026]	[0.028]	[0.025]
Peer	-0.129^{***}	-0.128^{***}	0.051	-0.044^{*}
	[0.024]	[0.026]	[0.031]	[0.024]
Cognition	-0.015	-0.079^{***}	0.011	0.018
	[0.030]	[0.031]	[0.035]	[0.030]
Family SES	0.086^{***}	-0.035	-0.011	0.038
	[0.026]	[0.026]	[0.032]	[0.024]
Background controls	X	X	X	X
Mean of Dep. Var.	0	0	0	0
N individual	2,637	3,377	1,847	2,876

Notes: Data from BCS70. Each column reports measurement-error-corrected estimates from regressions of age-16 behaviours on standardized socio-emotional skills, cognition, and family SES. Sport competition is a standardized measure of the number of sports in which teens represented their school/club over the past year. Teen socialization is a standardized predicted score obtained from our dedicated measurement system, see equation (1). Attitudes to Business is a standardized measure of the deviation of the interest for business with respect to mean occupational interest. Mental health is the standardized Malaise score (reverse-coded). All specifications include controls for: number of siblings, dummies for first child, no dad at birth, teenage mother, gender, and region at birth fixed effects. Standard errors in brackets are estimated from 250 bootstrap replications: ***p < .01, **p < .05, *p < .10.

The third column of Table 6 looks at determinants of occupational interests, introduced briefly in Section 3 and discussed in further depth in Appendix A. These data capture attitudes across six areas of work: business; practical; living; commerce; caring, and art. We focus here on interests in business, which the literature has found to be important (see e.g. Wiswall & Zafar, 2015) and which we expect to be most related to later earnings. Table 6 shows that relative interests in business are strongly negatively affected by problems of attention. It therefore seems that, not only do these problems of inattention negatively affect grades directly (Richardson et al., 2012), they also seem to affect relevant career aspirations which are nurtured at school.[27] Finally, the fourth column shows that mental health at age 16 is, not surprisingly, most strongly associated with emotional problems at age 10. We show gender differences in estimated effects in Table C.11.

We turn now to examining how these age-16 behaviours relate to later labour market outcomes, and how controlling for them, as possible mediating mechanisms, modifies the relationship between childhood skills and adult outcomes. This exercise, shown in Figure C.4, reveals two key findings. First, consistent with the social skills literature, sports competition has a positive and significant association with adult earnings, with a similar effect size (approximately 4%) as found by Deming (2017) for his participation measures.[28] Second, and quite strikingly, the estimated relationships between age-10 skills and earnings are little changed when these age-16 controls are included: point estimates on attention, conduct, emotional and peer problems shift by at most 0.015 log points (typically much less), and well within confidence intervals.[29]

Our overall conclusion is that age-16 measures signal important pathways from age-10 skills to adult outcomes, working in expected ways. However, they explain only a small share of the association between childhood skills and later labour-market outcomes. This suggests that age-16 behaviours are related but only partial expressions of the underlying socio-emotional skills measured at age 10. A fuller comparison with the social-skills literature (particularly Deming, 2017; Weinberger, 2014), including the absence of cognition–conduct complementarity in our age-10 measures discussed in Section 6.2, is given in Appendix D.3.

6.5 Health, Well-Being and Risky Behaviours

Our skill measures also predict a rich set of adult outcomes beyond the labour market. Figure 1 reports estimates from nine separate regressions of measures of well-being, mental health, and health-related behaviour on the age-10 skills and standard background controls.[30] The picture is one of sharply differentiated profiles across the four factors. This is most striking for conduct, which predicts both the highest rates of risky behaviour and, consistent with the labour market results, the highest job satisfaction.

In more detail, the top half shows relationships with health-related behaviours. Most noteworthy is the evidence of a negative association between internalising (emotion and peer) problems and these later behaviours. On the other hand we see a pronounced *positive* association of these later behaviours with conduct problems in childhood, which seem to raise the incidence of adult

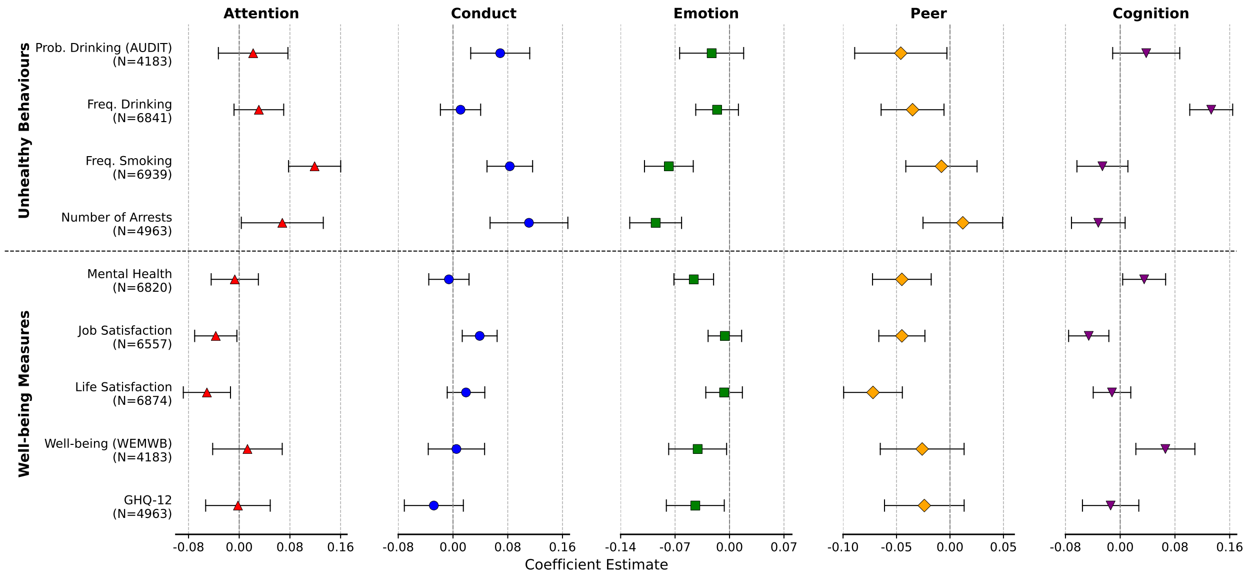
[27] For completeness, we show parallel regressions for attitudes to the other areas of work in Table C.12.

[28] In our benchmark results teen socialization has an insignificant effect on earnings. When controlling for years of schooling, its direct effect is positive. Results available on request.

[29] Table D.5 shows the symmetric result, showing that the estimated effects of age-16 behaviours are little changed when we remove controls for age-10 skills.

[30] In line with our baseline specification for labour market outcomes, schooling is omitted as a mediator. Behavioural skill coefficients are virtually unchanged when schooling is added.

Figure 1: Determinants of Adult Health and Well-being



Notes: Data from BCS70. The figure shows measurement-error-corrected estimates from regressions of standardised measures of well-being and unhealthy behaviours on socio-emotional skills, cognition, and family socio-economic status, constructed using our dedicated measurement system (see equation 1). Outcomes are ordered from top to bottom as follows. AUDIT: the Alcohol Use Disorders Identification Test, a 20-point score (2012, 2016). Frequency of drinking and smoking: Likert scales from "never" to "4 or more times per week" and "never" to "every day", respectively; drinking available in all years except 2008 and smoking in all years. Number of arrests: self-reported arrests before age 30 (2000). Mental health: the reversed Malaise score, based on 24 items in 1996 and 2000 and 9 items thereafter, standardised within each year. Job and life satisfaction: Likert scales ranging from "very dissatisfied" to "very satisfied" and "completely dissatisfied" to "completely satisfied", respectively; job satisfaction available in all years except 2004 and 2008 and life satisfaction in all years except 2008. WEMWB: the Warwick–Edinburgh Mental Well-Being Scale, a 70-point score (2012, 2016). GHQ-12: a 12-item general health score (2000). All specifications include controls for gender, gender-by-year, number of siblings, dummies for first child, no father at birth, teenage mother, and year and region fixed effects. Standard errors are estimated from 250 bootstrap replications clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

drinking, smoking and arrests. These findings add to the overall picture of varied outcomes for children displaying these characteristics.[31] Aside from these relationships, the figure shows an association of cognition with frequency of drinking, which possibly reflects the complex role of drinking in social behaviour as well as in stress relief. Similarly attention problems are associated with higher smoking, in line with the self-medication hypothesis (Wilens et al., 2007).

The next five rows show various measures of well-being and mental health. In line with Goodman et al. (2015) we find negative associations with internalising behaviours, with stronger associations for peer problems with satisfaction measures and relatively stronger associations for emotional problems with mental health. Attention problems are associated with significantly lower satisfaction measures, but with no noticeable effect on mental health as clinically measured. Interestingly, higher cognition is associated with worse job satisfaction. Most noteworthy however, is

[31] Webbink et al. (2012) show that conduct disorder has a negative effect on schooling and violent and criminal behaviour later in life. The study leverages within-family differences in self-reported behaviours from a sample of Australian twins.

the effect of conduct problems: worse conduct in childhood is associated with *higher* satisfaction in work. This result is clearly consistent with our results on the labour market.

7 Inequality and Intergenerational Mobility

7.1 Skill Gaps by Family Socio-Economic Status

A central question in economics concerns the intergenerational transmission of socio-economic advantage. Since at least Cunha and Heckman (2007), researchers have recognized that childhood skills, including non-cognitive abilities, correlate strongly with parental socio-economic status, suggesting that skill development represents a key transmission mechanism. We now examine how each of our skill measures varies by family SES and quantify their contribution to gaps in later life outcomes.

Table F.1 shows differences in levels of our standardized skill measures by family SES. The measures of SES used here are (i) a binary split of being above or below median in our latent family SES factor (Panel A) and (ii) whether or not either parent holds a college degree (Panel B). As expected, we see higher levels of attention, conduct, emotion and peer problems and a lower level of cognitive skills in the more disadvantaged group. The table also presents results for a test of first-order stochastic dominance, using the test proposed by Barrett and Donald (2003). The p-values reported in the final column show strongly that the distributions are consistent with the more advantaged group first-order stochastically dominating the lower group on all skill measures.[32] Appendix Figure F.1 shows the underlying densities for the split by family SES, revealing a particularly large separation for cognition and attention.

We also investigate the extent to which returns to skills differ by family SES. Table F.2 shows coefficients obtained by running separate regressions of schooling and earnings on our measures of childhood skills. As documented, there is no systematic evidence that the coefficients are statistically different by SES. The main exception is for cognitive skills, where we see higher returns for high-SES children, especially in terms of schooling. In Panel B we also see evidence that attention problems are more damaging for children of higher educated parents, but the coefficient for this group is not very precisely estimated. Overall, we conclude that returns to socio-emotional skills do not differ by family SES, which in turns implies that these skills matter for the transmission of socio-economic advantage mainly because of differences in endowments.

[32] The procedure tests the null that the high-SES CDF lies everywhere below the low-SES CDF (Lee & Whang, 2023). The zeros in the third column of Table F.1 indicate that the empirical high-SES CDF lies completely below the low-SES CDF across the entire distribution; small positive values reflect minor violations consistent with sampling variability. See Table notes for more details.

7.2 Decomposing Intergenerational Transmission of Economic Advantage

With these pieces of evidence in hand, we quantify the role of skills in the transmission of advantage using a Kitagawa-Oaxaca-Blinder decomposition. For completeness, and to explain how we account for noise in our skill measures, Appendix F formalizes the decomposition in our context.

The main results of the decomposition are shown in Table 7. The top rows of each panel show differences in outcomes (schooling and log earnings) by measure of SES. Children in the top half of the family SES distribution obtain 1.4 more years of schooling on average and earn 22% (or 20 log points) more than those in the bottom half, while children of those with degrees obtain 2.2 more years of schooling and earn around 27% (24 log points) more. The fact that gaps are largest by parental education is consistent with parental skills *per se* being more important than other resources in determining children's outcomes (see for example Heckman & Mosso, 2014).

The highlighted rows then show that across these two outcomes and two SES groupings, the explained part (i.e. the part attributable to differences in endowments) accounts for about half of the overall gap in outcomes. Among childhood skills, the contribution of cognitive skills is by far the most significant, accounting for around 45% of the explained earnings gap and close to 90% of the explained schooling gap. However, even after taking into account cognition, differences in socio-emotional skills remain statistically important. In terms of magnitude, differences in attention skills explain about 9% of the explained gap in schooling, and little more than 11% of the explained gap in log earnings. Similar results are obtained when we represent SES differences according to parental education. The analysis on earnings also shows an interesting feature reflecting our previous results. The SES gap in conduct problems partly *compensates* for differences in other childhood skills, offsetting around 6% of the explained earnings gap. This is because children from lower SES background have higher conduct problems, and conduct problems have been shown to be positively related to earnings.

Overall, these results are broadly consistent with those in Blanden et al. (2007) and Bolt et al. (2021), who also find that socio-emotional skills have a smaller quantitative impact than cognitive skills on the intergenerational transmission of (dis)advantage. However, by using a finer representation of socio-emotional skills, and specifically one that separates attention from conduct problems, our analysis brings additional insights. First, we see that SES differences in conduct problems are potentially equality-enhancing. Secondly, among all the socio-emotional skills considered, attention skills are found to be the most relevant component of the explained portion of SES differences in schooling and earnings. This suggests that interventions aimed at fostering skills such as perseverance and self-regulation can reduce SES inequalities later on in life as well as improve individual outcomes.[33]

[33] Studies by Alan et al. (2019) and Sorrenti et al. (2025) provide strong evidence that it is possible to increase these skills through school-based interventions.

Table 7: KOB Decomposition of Schooling & Earnings Gaps by Family SES

	Schooling		Earnings	
	Explained	Overall	Explained	Overall
Panel A: Top vs. Bottom Half Family SES				
Bottom-half Family SES		11.594***		7.077***
		[0.032]		[0.009]
Top-half Family SES		13.002***		7.275***
		[0.043]		[0.010]
Difference (Top − Bottom)	**-0.708*****	**-1.408*****	**-0.103*****	**-0.198*****
	[0.030]	**[0.054]**	**[0.009]**	**[0.013]**
Contribution by skill endowments:				
Attention	-0.061***		-0.011***	
	[0.016]		[0.003]	
Conduct	-0.013*		0.006***	
	[0.007]		[0.001]	
Emotion	0.010		-0.004***	
	[0.006]		[0.001]	
Peer	0.003		-0.002**	
	[0.005]		[0.001]	
Cognition	-0.628***		-0.045***	
	[0.032]		[0.004]	
Panel B: College Degree vs. No Degree				
No Degree		11.940***		7.136***
		[0.026]		[0.007]
Degree		14.104***		7.376***
		[0.068]		[0.016]
Difference (Degree − No Degree)	**-0.869*****	**-2.163*****	**-0.124*****	**-0.240*****
	[0.038]	**[0.071]**	**[0.013]**	**[0.018]**
Contribution by skill endowments:				
Attention	-0.074***		-0.010***	
	[0.020]		[0.003]	
Conduct	-0.014*		0.006***	
	[0.008]		[0.002]	
Emotion	0.012		-0.005***	
	[0.007]		[0.002]	
Peer	0.002		-0.002**	
	[0.005]		[0.001]	
Cognition	-0.761***		-0.059***	
	[0.041]		[0.005]	
Backg. Controls	X	X	X	X
N individuals	6,952	6,952	6,952	6,952
N individual-years			23,451	23,451

Notes: Data from BCS70. The table reports Kitagawa–Oaxaca–Blinder decompositions of gaps in years of schooling and log monthly earnings between children from different socioeconomic backgrounds. Panel A compares the bottom and top halves of the family SES distribution; Panel B compares children whose parents hold a college degree with those whose parents have no qualifications. "Overall" is the total gap between groups; "Explained" is the component attributable to differences in mean cognitive and socio-emotional endowments between groups. All models include controls for gender, birth order, number of siblings, no father at birth, teenage mother, and region fixed effects. Earnings regressions additionally include year fixed effects and gender-by-year interactions. Standard errors in square brackets are clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

8 Summary and Conclusions

In this paper we investigate the relationship between early measures of socio-emotional skills and later economic outcomes using data from the 1970 British Cohort Study. Using a measurement model based on teacher assessments of childhood development, we identify a four-factor representation of socio-emotional skills and explore its association with various adult outcomes. Our findings reveal that attention, peer and emotional problems are *negatively* related to earnings, while — contrary to conventional hypotheses — conduct problems are predictive of *positive* labour market outcomes.

The findings add to the existing literature on the economic returns to socio-emotional skills in four distinct ways. First, we present a new and methodologically robust analysis of childhood socio-emotional skills using teacher assessments. By showing that the resulting four-factor structure maps closely into the domains of the Strengths and Difficulties Questionnaire (SDQ), we also offer the first evidence on how SDQ-aligned childhood measures relate to later economic outcomes. Second, we show that while behavioural problems are generally associated with worse outcomes later in life, problems of conduct are associated with beneficial outcomes in the labour market, and a separate fixed-effects analysis conducted on similar data supports a causal interpretation. Third, by analysing age-16 behaviours, we show that competitive sport participation, teen socialization, career interests, and mental health are strongly associated with childhood socio-emotional skills and predictive of adult outcomes, but do not mediate much of the effect of early skills on earnings and schooling. Fourth, we present new evidence on how childhood skills contribute to the transmission of socio-economic advantage, finding that problems of attention are particularly relevant in exacerbating socio-economic inequalities. Beyond labour market outcomes, the four factors also generate sharply differentiated profiles across health, well-being, and risky behaviours: most strikingly, conduct problems predict both higher earnings and job satisfaction alongside elevated rates of drinking, smoking, and arrests.

Several policy-relevant implications emerge. The result that child socio-emotional skills are predictive of a range of adult outcomes, even conditional on extensive confounders and mediators, provides strong support for policies and interventions focused on the development of these skills in the early years, and for integrating socio-emotional learning into school curricula. Although this need is already recognised in the UK educational context, no uniform approach has emerged as yet (see Clarke et al., 2015 and Donnelly et al., 2020 for recent reviews of UK initiatives). The positive association between conduct problems and labour market outcomes further suggests a need to shift from punitive approaches towards understanding and addressing the underlying causes of disruptive behaviour. Recent evaluations of restorative justice practices, designed to address behavioural issues through constructive engagement rather than sanctions, demonstrate improved outcomes for children with conduct issues in several schools in Chicago (Adukia et al., 2025). However, further research is needed to understand whether such policies also improve the broader learning environment (including teacher well-being), which can be negatively affected by the presence of children with behavioural problems (Lavy & Yancu, 2025).

Several limitations are worth mentioning. First, and perhaps most importantly, the study relies on childhood skills data collected in the early 1980s, and teacher assessments and questionnaire wording reflect the educational norms of that period. This factor could affect the applicability of the results to more recent settings. That said, to the extent that the underlying skills are comparable, the literature on rising returns to social skills in an era of automation and AI (Deming, 2017; Edin et al., 2022) suggests that our estimates, if anything, likely understate the returns today's children can expect. Second, as discussed above, the data suffer from attrition and, while we conduct some checks to assess the effect on our results, this is an issue we cannot fully address. Finally, like many other studies in this literature, our analysis is observational. Although we employ several approaches to strengthen a causal interpretation, including sibling fixed effects to account for shared family background, other aspects such as peer influences or within-family dynamics may confound the observed associations.

References

Acemoglu, D., & Restrepo, P. (2018): The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, *108*(6), 1488–1542.

Acemoglu, D., & Restrepo, P. (2019): Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, *33*(2), 3–30.

Adukia, A., Feigenberg, B., & Momeni, F. (2025): From retributive to restorative: An alternative approach to justice. *American Economic Review*, *115*, 2722–54.

Alan, S., Baysan, C., Gumren, M., & Kubilay, E. (2021): Building social cohesion in ethnically mixed schools: An intervention on perspective taking. *The Quarterly Journal of Economics*, *136*(4), 2147–2194.

Alan, S., Boneva, T., & Ertac, S. (2019): Ever failed, try again, succeed better: Results from a randomized educational intervention on grit. *The Quarterly Journal of Economics*, *134*(3), 1121–1162.

Alderotti, G., Rapallini, C., & Traverso, S. (2023): The big five personality traits and earnings: A meta-analysis. *Journal of Economic Psychology*, *94*, 102570.

Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011): Personality psychology and economics. In *Handbook of the Economics of Education* (pp. 1–181, Vol. 4). Elsevier.

Attanasio, O., Blundell, R., Conti, G., & Mason, G. (2020): Inequality in socio-emotional skills: A cross-cohort comparison. *Journal of Public Economics*, *191*, 104–171.

Attanasio, O., De Paula, Á., & Toppeta, A. (2020): *The persistence of socio-emotional skills: Life cycle and intergenerational evidence*, National Bureau of Economic Research, 27823.

Autor, D. (2022): *The labor market impacts of technological change: From unbridled enthusiasm to qualified optimism to vast uncertainty*, National Bureau of Economic Research, 30074.

Barrett, G. F., & Donald, S. G. (2003): Consistent tests for stochastic dominance. *Econometrica*, *71*(1), 71–104.

Belfi, B., & Borghans, L. (2025): The importance of socio-emotional skills for multiple life outcomes and the role of education. In *Handbook of education and work* (pp. 75–104). Edward Elgar Publishing.

Bell, B. A., Taylor, S. G., Roberts, A. M., Shi, D., Burgess, K., Hough, C., & Flory, K. (2024): Factor structure of the teacher strengths and difficulties questionnaire in a large community-based sample: An investigation of alternative measurement models. *Assessment*, *31*(2), 291–303.

Bentler, P. M. (1990): Comparative fit indexes in structural models. *Psychological bulletin*, *107*(2), 238.

Blanden, J., Gregg, P., & Macmillan, L. (2007): Accounting for intergenerational income persistence: Noncognitive skills, ability and education. *The Economic Journal*, *117*(519), C43–C60.

Bolt, U., French, E., Maccuish, J. H., & O'Dea, C. (2021): *The intergenerational elasticity of earnings: Exploring the mechanisms*, CEPR Discussion Paper No. DP15975.

Borghans, L., Duckworth, A. L., Heckman, J. J., & Ter Weel, B. (2008): The economics and psychology of personality traits. *Journal of Human Resources*, *43*(4), 972–1059.

Borghans, L., Ter Weel, B., & Weinberg, B. A. (2014): People skills and the labor-market outcomes of underrepresented groups. *ILR Review*, *67*(2), 287–334.

Bound, J., & Solon, G. (1999): Double trouble: On the value of twins-based estimation of the return to schooling. *Economics of Education Review*, *18*(2), 169–182.

Bridger-Staatz, C., Kelly, Y., Lacey, R. E., & Hardy, R. (2024): Investigating the factorial structure and measurement invariance of the parent-reported strengths and difficulties questionnaire at 11 years of age from the uk millennium cohort study. *European Child & Adolescent Psychiatry*, *33*(1), 255–266.

Brodman, K., Erdmann, A. J., Lorge, I., Wolff, H. G., & Broadbent, T. H. (1951): The cornell medical index-health questionnaire: Ii. as a diagnostic instrument. *Journal of the American Medical Association*, *145*(3), 152–157.

Buss, D. M., & Shackelford, T. K. (1997): Human aggression in evolutionary psychological perspective. *Clinical Psychology Review*, *17*(6), 605–619.

Butler, N., Despotidou, S., & Shepherd, P. (1980): *British cohort study (bcs70) ten-year follow-up: A guide to the bcs70 10-year data)* (tech. rep.). Social Statistics Research Unit, City University.

Campbell, D. T., & Fiske, D. W. (1959): Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105.

Chazan, M. (1976): *Studies of infant school children*. B. Blackwell.

Chazan, M. (1982): Disadvantage, behaviour problems and reading difficulties. *AEP (Association of Educational Psychologists) Journal*.

Chazan, M., & Threlfall, S. (1978): Behavior problems in infant school children in deprived areas. *The Child in His Family: Vulnerable Children, New York: John Wiley & Sons*.

Clarke, A. M., Morreale, S., Field, C.-A., Hussein, Y., & Barry, M. M. (2015): *What works in enhancing social and emotional skills development during childhood and adolescence. a review of the evidence on the effectiveness of school-based and out-of-school programmes in the uk.* (tech. rep.). The World Health Organization Collaborating Centre for Health Promotion Research, National University of Ireland Galway.

Closs, S. (1978): An algorithm for integrating preference and like–dislike responses to the same test items. *Journal of Occupational Psychology*, *51*(2), 147–154.

Conners, C. K. (1969): A teacher rating scale for use in drug studies with children. *American Journal of Psychiatry*, (126), 884–888.

Conners, C. K., Sitarenios, G., Parker, J. D., & Epstein, J. N. (1998): Revision and restandardization of the conners teacher rating scale (ctrs-r): Factor structure, reliability, and criterion validity. *Journal of abnormal child psychology*, *26*, 279–291.

Conti, G., Frühwirth-Schnatter, S., Heckman, J. J., & Piatek, R. (2014): Bayesian exploratory factor analysis. *Journal of Econometrics*, *183*(1), 31–57.

Cunha, F., & Heckman, J. (2007): The technology of skill formation. *American economic review*, *97*(2), 31–47.

Currie, J., & Almond, D. (2011): Human capital development before age five. In *Handbook of labor economics* (pp. 1315–1486, Vol. 4). Elsevier.

Darling-Hammond, S., & Fronius, T. (2022): Restorative practices in schools. *Handbook of classroom management*, 54–73.

Del Bono, E., Kinsler, J., & Pavan, R. (2026): Skill formation and the trouble with child noncognitive skill measures. *Quantitative Economics*, *17*(1), 135–172.

Deming, D. J. (2017): The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, *132*(4), 1593–1640.

Dickerson, A., & Morris, D. (2019): The changing demand for skills in the uk. *Centre for Vocational Education Research, Research Paper*, *20*.

Donnelly, M., Brown, C., Batlle, I. C., & Sandoval-Hernández, A. (2020): *Developing social and emotional skills: Education policy and practice in the uk home nations* (tech. rep.). Nesta Reports.

Edin, P.-A., Fredriksson, P., Nybom, M., & Öckert, B. (2022): The rising return to noncognitive skill. *American Economic Journal: Applied Economics*, *14*(2), 78–100.

Ehrmantraut, L., Pinger, P., & Stans, R. (2020): *The expected (signaling) value of higher education*, CESifo Working Paper, 8589.

Feng, S., Han, Y., Heckman, J. J., & Kautz, T. (2022): Comparing the reliability and predictive power of child, teacher, and guardian reports of noncognitive skills. *Proceedings of the National Academy of Sciences*, *119*(6), e2113992119.

Fitzsimons, E., & Vera-Hernández, M. (2022): Breastfeeding and child development. *American Economic Journal: Applied Economics*, *14*(3), 329–366.

Fletcher, J. M. (2013): The effects of personality traits on adult labor market outcomes: Evidence from siblings. *Journal of Economic Behavior & Organization*, *89*, 122–135.

Fortin, N., Lemieux, T., & Firpo, S. (2011): Decomposition methods in economics. In *Handbook of labor economics* (pp. 1–102, Vol. 4). Elsevier.

Gammage, P. (1975): *Socialisation, schooling and locus of control* [Doctoral dissertation, Bristol University].

Goodman, A., Joshi, H., Nasim, B., & Tyler, C. (2015): Social and emotional skills in childhood and their long-term effects on adult life. *London: Institute of Education*.

Goodman, R. (1997): The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586.

Goodman, R., Ford, T., Simmons, H., Gatward, R., & Meltzer, H. (2000): Using the strengths and difficulties questionnaire (sdq) to screen for child psychiatric disorders in a community sample. *The British journal of psychiatry*, *177*(6), 534–539.

Goodman, R., Lamping, D. L., & Ploubidis, G. B. (2010): When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the strengths and difficulties questionnaire (sdq): Data from british parents, teachers and children. *Journal of Abnormal Child Psychology*, *38*, 1179–1191.

Hall, C. L., Guo, B., Valentine, A. Z., Groom, M. J., Daley, D., Sayal, K., & Hollis, C. (2019): The validity of the strengths and difficulties questionnaire (sdq) for children with adhd symptoms. *PloS one*, *14*(6), e0218518.

Heckman, J., Pinto, R., & 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., Stixrud, J., & Urzua, S. (2006): The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, *24*(3), 411–482.

Heckman, J. J., & Mosso, S. (2014): The economics of human development and social mobility. *Annu. Rev. Econ.*, *6*(1), 689–733.

Holmlund, H., Lindahl, M., & Plug, E. (2011): The causal effect of parents' schooling on children's schooling: A comparison of estimation methods. *Journal of economic literature*, *49*(3), 615–651.

Hopkins, B. (2023): *The restorative classroom: Using restorative approaches to foster effective learning*. Routledge.

Horn, J. (1965): A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185.

Izadi, R., & Tuhkuri, J. (2024): Evolving returns to personality. *Journal of Labor Economics*.

Jin, W. (2022): *Occupational polarisation and endogenous task-biased technical change* (tech. rep.). University College London.

Johnston, D., Propper, C., Pudney, S., & Shields, M. (2014): Child mental health and educational attainment: Multiple observers and the measurement error problem. *Journal of Applied Econometrics*, *29*, 880–900.

Jokela, M., Pekkarinen, T., Sarvimäki, M., Terviö, M., & Uusitalo, R. (2017): Secular rise in economically valuable personality traits. *Proceedings of the National Academy of Sciences*, *114*(25), 6527–6532.

Joreskog, K. (1971): Statistical analysis of sets of congeneric tests. *Psychometrika*, *36*, 109–133.

Jöreskog, K. G., & Sörbom, D. (1986): Lisrel vi: Analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods. *Scientific Software, Inc Mooresville, IND*, *4*.

Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., & Borghans, L. (2014): *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success*, National Bureau of Economic Research, 20749.

Lavy, V., & Yancu, A. (2025): *Violent peers at school: Impacts and mechanisms* (tech. rep.). National Bureau of Economic Research.

Lawrence, D. (1981): The development of a self-esteem questionnaire. *British Journal of Educational Psychology*, *51*(2), 245–251.

Lee, K., & Whang, Y.-J. (2023): Pysdtest: A python/stata package for stochastic dominance tests. *arXiv preprint arXiv:2307.10694*.

Levine, R., & Rubinstein, Y. (2017): Smart and illicit: Who becomes an entrepreneur and do they earn more? *The Quarterly Journal of Economics*, *132*(2), 963–1018.

Lewis, G. J., Haworth, C. M., & Plomin, R. (2014): Identical genetic influences underpin behavior problems in adolescence and basic traits of personality. *Journal of Child Psychology and Psychiatry*, *55*(8), 865–875.

Lim, S., & Jahng, S. (2019): Determining the number of factors using parallel analysis and its recent variants. *Psychological Methods*, *24*(4), 452–467.

Lindqvist, E., & Vestman, R. (2011): The labor market returns to cognitive and noncognitive ability: Evidence from the swedish enlistment. *American Economic Journal: Applied Economics*, *3*(1), 101–128.

Lise, J., & Postel-Vinay, F. (2020): Multidimensional skills, sorting, and human capital accumulation. *American Economic Review*, *110*(8), 2328–2376.

McCrae, R. R., & Costa, P. T. (1999): A five-factor theory of personality. *Handbook of personality: Theory and research*, *2*(1999), 139–153.

Muris, P., Meesters, C., & Diederen, R. (2005): Psychometric properties of the big five questionnaire for children (bfq-c) in a dutch sample of young adolescents. *Personality and individual differences*, *38*(8), 1757–1769.

Niclasen, J., Teasdale, T. W., Andersen, A.-M. N., Skovgaard, A. M., Elberling, H., & Obel, C. (2012): Psychometric properties of the danish strength and difficulties questionnaire: The SDQ assessed for more than 70,000 raters in four different cohorts. *PloS One*, *7*(2), e32025.

Nicoletti, C., & Vidiella-Martin, J. (2025): ADHD, school performance, and economic outcomes. *Oxford Research Encyclopedia of Economics and Finance*.

Papageorge, N. W., Ronda, V., & Zheng, Y. (2019): *The economic value of breaking bad: Misbehavior, schooling and the labor market*, National Bureau of Economic Research, 25602.

Prevoo, T., & ter Weel, B. (2015): The importance of early conscientiousness for socio-economic outcomes: Evidence from the british cohort study. *Oxford Economic Papers*, *67*(4), 918–948.

Reyes, J. W. (2015): Lead exposure and behavior: Effects on antisocial and risky behavior among children and adolescents. *Economic Inquiry*, *53*(3), 1580–1605.

Richardson, M., Abraham, C., & Bond, R. (2012): Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353.

Rodgers, J. L., Beasley, W. H., Bard, D. E., Meredith, K. M., D. Hunter, M., Johnson, A. B., Buster, M., Li, C., May, K. O., Mason Garrison, S., et al. (2016): The nlsy kinship links: Using the nlsy79 and nlsy-children data to conduct genetically-informed and family-oriented research. *Behavior Genetics*, *46*(4), 538–551.

Rutter, M. (1967): A children's behaviour questionnaire for completion by teachers: Preliminary findings. *Child Psychology & Psychiatry & Allied Disciplines*.

Segal, C. (2013): Misbehavior, education, and labor market outcomes. *Journal of the European Economic Association*, *11*(4), 743–779.

Sjölander, A., Frisell, T., & Öberg, S. (2022): Sibling comparison studies. *Annual Review of Statistics and Its Application*, *9*(1), 71–94.

Slobodskaya, H. R. (2007): The associations among the big five, behavioural inhibition and behavioural approach systems and child and adolescent adjustment in russia. *Personality and Individual Differences*, *43*(4), 913–924.

Sorrenti, G., Zölitz, U., Ribeaud, D., & Eisner, M. (2025): The causal impact of socio-emotional skills training on educational success. *Review of Economic Studies*, *92*(1), 506–552.

Steiger, J. H. (1990): Structural model evaluation and modification: An interval estimation approach. *Multivariate behavioral research*, *25*(2), 173–180.

Tucker, L. R., & Lewis, C. (1973): A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1–10.

Vicentini, G., Raccanello, D., & Burro, R. (2025): Self-report questionnaires to measure big five personality traits in children and adolescents: A systematic review. *Scandinavian Journal of Psychology*.

Volk, A. A., Dane, A. V., & Al-Jbouri, E. (2022): Is adolescent bullying an evolutionary adaptation? a 10-year review. *Educational Psychology Review*, *34*(4), 2351–2378.

Webbink, D., Vujić, S., Koning, P., & Martin, N. G. (2012): The effect of childhood conduct disorder on human capital. *Health economics*, *21*(8), 928–945.

Weinberger, C. J. (2014): The increasing complementarity between cognitive and social skills. *Review of Economics and Statistics*, *96*(4), 849–861.

Wilens, T. E., Adamson, J., Sgambati, S., Whitley, J., Santry, A., Monuteaux, M. C., & Biederman, J. (2007): Do individuals with adhd self-medicate with cigarettes and substances of abuse? results from a controlled family study of adhd. *The American Journal on Addictions*, *16*, 14–23.

Wiswall, M., & Zafar, B. (2015): Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, *82*(2), 791–824.

Zakszeski, B., & Rutherford, L. (2021): Mind the gap: A systematic review of research on restorative practices in schools. *School Psychology Review*, *50*(2-3), 371–387.

Appendix for:

The Economic Value of Childhood Socio-Emotional Skills

Emilia Del Bono, Ben Etheridge & Paul Garcia

A Further Details on the Data and Methodology

A.1 Socio-emotional Development Scales

Our analysis uses three development scales designed by psychologists that have become increasingly used in the economics literature.

The Rutter Child Behaviour Questionnaire (Rutter, 1967), was one of the first standardized measures of behaviour of school-aged children and introduced a systematic approach to assessing behavioural and emotional problems. It consists of two versions: the parent scale (Scale A) with 31 items, and the teacher scale (B) comprising 26 items. Items cover behavioural difficulties, emotional problems, and social relationships, rated on a three-point Likert scale (Doesn't Apply, Applies Somewhat, Certainly Applies). In the BCS70 implementation, teachers did not use the discrete response categories from the original instruments. Instead, they marked a continuous line for each item indicating their level of agreement, which survey analysts subsequently coded as integers ranging from 1 to 47. This continuous coding provides finer variation than the original scales and is the basis for our factor analysis.

The Conners Teacher Rating Scale (Conners, 1969) was specifically designed to assess attention deficit and hyperactivity symptoms, focussing on classroom behaviours. The original version contained 39 items rated on a four-point Likert scale (Not at All, Just a Little, Pretty Much, Very Much). More recent versions of this scale remain commonly used for assessing ADHD symptoms in educational settings, and demonstrate reliability across different populations and age groups (Conners et al., 1998).

The Swansea Assessment Battery was developed in the 1960s in consultation with Professor Maurice Chazan, a distinguished figure in educational psychology at the time (Butler et al., 1980). The questionnaire was designed to assess various behavioural and emotional aspects of children, but the full set of items was never disseminated. Professor Chazan's work focused on the identification and assessment of emotional and behavioural difficulties in children aged 7 to 11 (Chazan, 1976, 1982; Chazan & Threlfall, 1978).

The rationale for combining items from these three scales was to augment the Rutter questionnaire's broad assessment of problem behaviour with more specific items on hyperactivity and inattention from the Conners and Swansea instruments, thereby providing richer coverage of the behavioural domain.

Throughout, we compare the measures derived from these scales to the Strengths and Difficulties Questionnaire (SDQ, Goodman, 1997) which is a screening tool for assessing behaviour of school-aged children and adolescents. The SDQ consists of 25 items across five scales: emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behaviour. Each item uses a three-point Likert scale (Not True, Somewhat True, Certainly

True). Since its development, the SDQ has become one of the most widely used measures of child mental health and behaviour, with validated translations in over 40 languages, and is used in both research and clinical settings (see e.g. Goodman et al., 2000).[34]

A.2 Variable Construction

Our years of schooling variable is defined as total completed years in school by age 26. Accordingly the vast majority of measures are taken from the age-26 sweep. For those not appearing in this sweep we take the first subsequent report in following sweeps. In practice, years of schooling is calculated from the highest qualification attained. For example, individuals leaving school at age 16 without completing O-levels are assigned 10 years, while those completing O-levels are assigned 11 years. Holders of bachelors degrees are assigned 16 years. In cases where respondents do not report a qualification level, we use the reported age of leaving education instead, where available.

Our earnings variable is usual take-home pay excluding bonuses, defined for employees only. Survey respondents report the relevant time period for which their earnings are received, which we use to convert to a monthly frequency. To address extreme values, we winsorize the top and bottom 3% of values within sweeps.

Occupational codes from waves 1996 to 2012 are harmonized to the UK SOC 2010 three-digit classification (around 90 occupations) using standard crosswalk tables available from ONS.[35] Occupational characteristics from US O*NET data are matched to UK SOC 2010 codes using ISCO-08 following a standard procedure in the literature (see e.g. Dickerson & Morris, 2019; Jin, 2022).

For hours, we use usual hours worked per week (excluding meal breaks). Similarly to earnings we winsorize the top and bottom 3% within sweeps. Self-employed individuals, who typically do not report hours, are excluded from the sample. Wages for employees are computed as the ratio of earnings to hours.

Family socio-economic status is derived from a factorization of four variables. The first variable is family income, categorized into eight bands. The second variable captures parent qualifications. For this we first use the father's qualification level. If father's qualifications are not reported, we use the qualification level of the mother. The third variable is the survey's own measure of socio-economic status, given by the broad occupation group of the father or, if the father is unemployed or absent, the occupation of the mother. Last, we include a binary variable indicating whether either parent holds a managerial position. In cases where one of these variables is missing (e.g., occupational status when both parents are unemployed), we impute values as discussed below. All of these variables are reported in the age-10 sweep.

[34]See the SDQ's website: https://www.sdqinfo.org/a0.html, accessed on 23/02/2025.

[35]The data are publicly available at SOC 2010.

Other control variables are obtained directly from the dataset. These include categorical variables indicating the presence of the father in the household at the time of birth, whether the mother was a teenager at the time of birth, and the region of birth, all collected in the original age-0 sweep. Additionally, we consider the number of siblings and birth order as reported in the age-10 sweep.

We also incorporate items related to socialization at age 16, as documented in Table B.7. For assessing mental health, we consider the total malaise score reported in the survey at age 16, which encompasses 22 items covering aspects such as quality of sleep, feelings of misery, and occurrence of headaches or backaches. This score ranges from 0 to 44 and is standardized for our analysis. We also compute a measure of participation in competitive sports as detailed in Section 3.

A.3 Data on Occupational Interests

As mentioned in Section 3 we use data on occupational interests from the BCS survey collected in the age-16 sweep, in 1986. These data were archived un-processed for many years and only released in 2016. They come from a computer-assisted survey in which respondents were presented with a list of 30 pairwise options for jobs they would like to perform. These options were arranged in six menus which were designed to depend on the respondent's hypothetical skill level in the labour market, and which the respondent was allowed to select depending on their anticipated final schooling level. For example the first menu asks respondents to choose between 'Repair holes in roads' and 'Lift potatoes from fields', while the sixth menu asks respondents to choose between 'Do research on new ways of producing energy' and 'Study the causes of diseases'. In their chosen menu, respondents were forced to choose a preferred outcome, but also asked to rate each outcome with an intensity of their interest. The raw data were then transformed using the Closs algorithm (Closs, 1978), which provides individual-level scores on the overall interest in each of six career tracks: business occupations; practical; living; communication; art; and caring occupations.

In our analysis we use these transformed scores. As presented in Tables 6, C.11 and C.12 we use the interest score for the relevant career track as the outcome variable. To control for overall enthusiasm for work we control for the total score across all career tracks, which varies noticeably across individuals and skill level. Accordingly, our results should be interpreted as capturing *relative interest* in a specific career, where some children have more (or less) interest in higher-paying occupations relative to lower paying ones.

A.4 Imputation

Our statistical analysis faces challenges of attrition and non-response, which vary significantly across sweeps. We deal with missing items by using a random forest imputation algorithm under the assumption of missing at random.[36] Acknowledging that this is a strong assumption, we im-

[36]We compare performance among different types of imputation methods such as median/mode/mean, multiple imputation (predictive mean matching/polynomial regression), and K-nearest neighbours (KNN, euclidean/hamming

Table A.1: Summary Statistics of Main Variables

Variables	N	Mean	SD	Min	Max
Year	23451	2005.69	6.52	1996	2016
Time-Varying					
Log Monthly Earnings	23451	7.17	0.61	5.18	8.55
Log Hourly Wages	23451	2.16	0.48	-0.29	4.52
Log Monthly Hours	23451	5.02	0.35	3.95	5.58
London	23451	0.08	0.27	0	1
Fixed					
Years of Schooling	6952	12.26	2.31	0	17
Female	6952	0.51	0.50	0	1
Number siblings	6952	1.58	1.04	0	8
First child	6952	0.42	0.49	0	1
Dad not present (age 0)	6952	0.04	0.19	0	1
Teenage mother (age 0)	6952	0.08	0.28	0	1
Standardized variables					
Attention	6952	0	1	-1.47	3.10
Conduct	6952	0	1	-1.07	4.86
Emotion	6952	0	1	-1.60	2.85
Peer	6952	0	1	-1.82	3.28
Cognition	6952	0	1	-3.99	2.69
Family SES	6952	0	1	-2.02	2.64
Age-16 behaviours					
Sport Competition	2637	0	1	-0.61	4.79
Teen Socialization	3377	0	1	-2.33	2.12
Attitudes to Business	1847	0	1	-2.89	2.89
Mental Health	2876	0	1	-5.95	1.73

Notes: Data from BCS70. The table shows summary statistics for key variables of this study. Labor market information on earnings, wages, and hours, is estimated using information on 6 waves from 1996 to 2016 for the main sample (6952 individuals with complete information at age 10). Years of Schooling are estimated by using information on the highest academic qualification achieved and the age at which an individual leaves full-time education as reported in wave 1996. Missing years of schooling for non-participants in 1996 are completed using information from subsequent waves. Socio-demographic characteristics, factor scores for age-10 skills, and age-16 behaviours (e.g. business orientation, mental health, etc.) are obtained from waves 1970, 1980 and 1986. See Section 4 for detailed information on the construction of skills.

pute missing observations only for individuals with at most one missing item within each scale. Similarly, we impute missing test scores only for children with at most one missing cognitive test out of the eight available.

Table A.3 displays the number of tests and scales we use in our factor analysis and shows the number of individuals with complete cases, missing values and imputations performed for each scale/test. Although 14,870 individuals were interviewed at age-10, we have teacher responses to the Child Behaviour Scale for only 65% of them. Response rates are higher for self-reported scales such as LAWSEQ and CARALOC, at approximately 80% of the achieved sample. Column 4 summarizes the number of individuals with one or more missing items (excluding those who did not answer any question within a scale). As discussed, we perform random forest imputations using information from individuals who answered all but one question within each scale. As shown in column 5, around 80% of individuals with missing observations satisfy this criterion. In total, across the behavioural scales approximately 11,000 individuals provide complete responses, and we recover around 360 additional item-level observations through imputation. For the cognitive test battery, approximately 4,300 individuals have complete scores, with a further 2,000 recovered through imputation. Overall, a reasonably full set of measures is obtained and used in the analysis.

At age 16, the number of individuals with missing observations is much larger. To derive our measure of teen socialization, we use two questionnaires at age 16: 'Friends and Outside World (Document H)' (five items) and 'Attitudinal Scale - At Leisure (Document C)' (seven items). The former provides information on social behaviour during term time and the latter during leisure time. There are two versions of the Attitudinal scales: i) one administered at school during term time and ii) one administered at home after individuals finish the school year. Given that both questionnaires were administered on different dates and locations, the proportion of individuals with complete information in one questionnaire and no information in the other is large. Thus, we impute all missing observations, irrespective of their number, conditional on the individual answering all the relevant questions in the other questionnaire. This allows us to recover approximately 80% of the individuals with missing information within either the Friends and Outside the World Questionnaire or the Attitudinal Scales.

A.5 Factor Model: Identification, Estimation, and Scoring

This subsection details the identification assumptions, the iterative exploratory factor analysis (EFA) procedure, the confirmatory factor analysis (CFA) and scoring method underlying the factor

distance method) according to the characteristics of the variables (continuous/categorical). We observe that a non-parametric imputation method such as iterated random forests outperforms parametric methods as we use a mixture of numerical and categorical variables throughout the study. Moreover, random forests are robust to noisy data as they build in feature selection, unlike other methods such as KNN. This characteristic eases our concerns about meaningless data biasing our predicted imputed values, given the large number of variables in our analysis.

Table A.2: Sample Statistics: Age-10 and -16 Variables

Age 10	**N observations**	**Age 16**	**N observations**
Number of obs.	14870	Number of obs.	11815
Cognition			
(BAS, Maths, Reading Tests, 8 items)			
Complete	11110		
Imputed	367		
Total	11477		
Child Development			
(Rutter, Conners, 42 items)			
Complete	9686		
Imputed	2300		
Total	11986		
Family SES		**Teen Socialization**	
(family income, parental qual., 4 items)		(Friends & Outside, Att. Scales, 12 items)	
Complete	11850	Complete	4318
Imputed	1479	Imputed	2016
Total	13329	Total	6334
Complete info. Age-10	9345	Complete info. Age-10 + Age-16 social	3998
Info. labour market attachment	11520		
Complete info. Age-10 + labour market attachment	**6952**	Complete info. Age-10 + Age-16 social + labour market attachment	**3377**

Notes: Data from BCS70. The table outlines the count of completed and imputed cases for all age-10 and age-16 variables used in this study. The main sample consists of individuals with full information after imputation on age-10 measures (including locus of control scale (CARALOC) and self-esteem scale (Lawrence), mainly used in robustness checks) and labor market attachment. The sample is significantly smaller when including age-16 variables due to high attrition levels in wave 1986. Sample sizes are detailed in each table throughout the paper.

system used throughout the paper. The measurement model is stated in equation (1) of Section 4; here we supply the technical detail omitted from the main text.

Identification: The measurement model leaves the number of factors and the assignment of items to factors to be determined empirically. To achieve identification, and as standard, we impose $Cov(\epsilon_{\omega,j}, \epsilon_{\omega,j'}) = Cov(\omega, \epsilon_{\omega,j}) = 0$ for items $j \neq j'$, and $Var(\omega) = 1$.[37]

Iterative Exploratory Factor Analysis: The first challenge is to determine the number of factors and the item assignment. We proceed with a dedicated exploratory factor analysis (EFA), following, for example, Heckman et al. (2013). To make the analysis manageable, we group the items into

[37] An interesting issue, that we do not pursue here, is the extent to which different teachers adopt the same scaling factor. If some teachers systematically record higher scores than others, for observationally-equivalent children, then this would introduce a correlation of measurement errors across items that we are not accounting for. We address this issue to a certain extent in Section 6.3 where we discuss possible correlation of teacher reporting style with confounding school characteristics. We consider a fuller investigation of this point a topic for future research.

Table A.3: Number of Individuals with Complete, Missing, and Imputed Values by Scale/Test

	Number of items [1]	Answered at least one [2]	Answered all [3]	Missing obsv. [4]	% of imputed missing obs [5]	Total OOB error [6]
Scales/Tests at Age 10						
Cognitive Tests	8	12863	11110	1753	20.94%	0.30
Child Behaviour Scale	42	12702	9686	3016	76.26%	0.59
Family SES	4	13695	11850	1845	80.16%	0.48
Self-esteem (Lawrence)	12	12662	11850	812	81.77%	0.40
Locus of control (CARALOC)	16	12612	11713	899	80.65%	0.43
Scales at Age 16						
Friends, Outside World and Attitudinal Scales	12	6590	4318	2272	88.73%	0.45

Notes: Data from BCS70. The table shows the number of items per scale and the number of individuals with complete and missing values. Column [4] shows the total number of individuals with missing observations. Column [5] shows the percentage of observations that were imputed from column [4], that is, those individuals who answered all questions/tests but one within each scale. We performed a non-parametric imputation using Random Forest. Column [6] shows an estimate of the out-of-bag (OOB) imputation error by scale.

three broad domains: cognitive and socio-emotional skills; family background; and teen (age-16) socialization.

Within each domain we follow an iterative procedure. We perform a first EFA using all items, and then successively remove items either with factor loadings below 0.4, or which load highly on multiple factors — specifically with a ratio of loadings greater than 75%. The number of factors is chosen using the Kaiser criterion, which implies keeping only those factors whose eigenvalues are greater than 1.[38] The process terminates at the round when no further items and no factors are dropped. The resulting system aligns each factor with a specific set of items, so that the empirical content of each factor is highly transparent.

Confirmatory Factor Analysis and Bartlett Scoring: Having determined the factor structure by EFA, we estimate the final system by confirmatory factor analysis (CFA), setting all cross-loadings to zero. We then compute factor scores at the individual level using Bartlett scores, which aggregate item scores $Z_{\omega,i,j}$ as follows:

$$\omega_i^S = \kappa_\omega \sum_j \frac{\lambda_{\omega,j}}{\sigma^2_{\omega,\epsilon,j}} Z_{\omega,i,j} \tag{3}$$

where $\sigma^2_{\omega,\epsilon,j}$ is the variance of the noise on the jth item, equalling $1-\lambda^2_{\omega,j}$, and $\kappa_\omega \equiv \left(\sum_j \frac{\lambda^2_{\omega,j}}{\sigma^2_{\omega,\epsilon,j}}\right)^{-1}$ is a factor-specific scaling factor. All components can be replaced with sample counterparts. Bartlett

[38] A factor with an eigenvalue greater than 1 explains more variance than any single standardized item.

scores can be construed as the coefficient from a weighted regression of item scores on the loading factors at the individual level, and minimize noise among scoring methods. The scoring formula (3) is used in the bias-correction derivation of the following Appendix subsection.

A.6 Bias Correction for the Analysis of Long-Term Outcomes

Section 4 and Appendix A.5 discussed the identification and estimation of our factor model, which follows closely the approach in Heckman et al. (2013) and Bolt et al. (2021). Section 4 also discussed how the measurement system is used to correct for measurement-error-induced attenuation in our analysis of long-term outcomes. Here we specify the precise formula used to do that.

As previously discussed we construct factor scores ω_i^S for each of our six factors ω (problems of: attention, conduct, emotion and peer relations, together with family socio-economic status and teen socialization). These factor scores can then be expressed as:

$$\omega_i^S = \omega_i + \eta_i \tag{4}$$

where $\eta_i = \kappa_\omega \sum_j \frac{\lambda_{\omega,j}}{\sigma^2_{\omega,\epsilon,j}} \epsilon_{\omega,i,j}$ is a weighted sum of item-level measurement errors, is orthogonal to the latent factor, and is assumed orthogonal to covariates of interest. The factor-specific coefficient is defined above as $\kappa_\omega \equiv \left(\sum_j \frac{\lambda^2_{\omega,j}}{\sigma^2_{\omega,\epsilon,j}} \right)^{-1}$.

For the model expressed in equation (2) in Section 4, running OLS with $\Omega_{i,\langle a \rangle}$ replaced with the observed scores is inconsistent because these scores are then correlated with the regression error term, which includes η_i. However, the form of the bias can be derived. We rewrite equation (2) as:

$$y_{i,t} = \Gamma \Omega^S_{i,\langle a \rangle} + \beta \mathbf{X}_{i,t} + v_{i,t} - \Gamma \bar{\eta}_i$$

where $\Omega^S_{i,\langle a \rangle}$ is the vector of observed scores, and $\bar{\eta}_i$ is the vector of measurement errors. Then, abbreviating notation, and as discussed by Heckman et al. (2013) and Bolt et al. (2021):

$$\text{plim} \begin{pmatrix} \hat{\Gamma} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} \text{Cov}\left(\Omega^S, \Omega^S\right) & \text{Cov}\left(\Omega^S, \mathbf{X}\right) \\ \text{Cov}\left(\Omega^S, \mathbf{X}\right) & \text{Cov}\left(\mathbf{X}, \mathbf{X}\right) \end{pmatrix}^{-1} \begin{pmatrix} \text{Cov}\left(\Omega, \Omega\right) & \text{Cov}\left(\Omega, \mathbf{X}\right) \\ \text{Cov}\left(\Omega, \mathbf{X}\right) & \text{Cov}\left(\mathbf{X}, \mathbf{X}\right) \end{pmatrix} \begin{pmatrix} \Gamma \\ \beta \end{pmatrix} \tag{5}$$

Using expression (4) above, and using the orthogonality conditions, we note that $\text{Cov}\left(\Omega, \mathbf{X}\right) = \text{Cov}\left(\Omega^S, \mathbf{X}\right)$, which is observed. Similarly $\text{Cov}\left(\Omega, \Omega\right)$ is equal to the observed covariance of scores, $\text{Cov}\left(\Omega^S, \Omega^S\right)$, except with diagonal elements replaced with 1s. Equation (5) thus allows us to identify Γ and β.

B Detailed Factorization Results

This section provides supporting detail on the age-10 factorization summarized in Section 5, covering the items dropped, the correspondence with the SDQ (Table B.1), validation exercises (gender-specific EFA, the Rutter-only factorization, the multitrait-multimethod analysis), auxiliary factorizations of mother reports and of locus-of-control and self-esteem items, confirmatory factor analysis loadings (Tables B.8 and B.9), and the correlation matrix of factor scores (Table B.10).

Item retention and dropped items. Following Heckman et al. (2013) and Bolt et al. (2021), we drop items with loadings below 0.4 as unstable, and items that cross-load too highly between factors. Examples of items failing the loading criterion are 'cannot negotiate child's behaviour' and 'face or body twitches'; examples failing the cross-loading criterion are 'child is squirmy and fidgety' and 'child confused with difficult tasks'. Cronbach's alpha is comfortably above the usual threshold of 0.7 for all factors.[39] As further validation of the number of factors retained, Figure B.1 shows a scree plot of the first six eigenvalues from the final EFA iteration, together with the adjusted eigenvalues from Horn's parallel analysis.

Relationship with the Strengths and Difficulties Questionnaire. Table B.1 details the correspondence between our items and the SDQ scales. The first column shows SDQ item abbreviations with their corresponding scales; columns two and three show the closest matching BCS70 items and the factor to which they are assigned. One minor point not discussed in the main text is that our peer factor includes 'child is not cooperative', which most closely matches the SDQ item 'volunteers to help others' from the pro-social scale. This item fits reliably into our peer problems factor.

Gender-specific factorization. To ensure a consistent set of items for all individuals, we perform the EFA pooling across men and women. As a check that our approach picks up the same factors across genders, Table B.3 shows the final iteration of a gender-specific EFA, reporting the full set of loadings across all factors for the final set of items. The loadings are very similar by gender, and internal consistency remains above 0.8 for all factors even in the smaller gender-specific samples.

Two-factor (Rutter-only) factorization. Table B.2 shows the loadings from the Rutter-only factorization referenced in the main text and their alignment with our preferred factorization. Items on the externalising factor mostly line up with our conduct factor, with one item ('squirmy and fidgety') more aligned with attention; items on the internalising factor line up with our emotional and peer factors. Three of the items retained here — 'squirmy and fidgety', 'face or body twitches', and 'unhappy or tearful' — are not retained in our primary analysis.

[39]Cronbach's alpha ranges from 0 to 1, capturing the average correlation of items within a factor across individuals, and is a standard measure of internal consistency of an underlying construct.

Multitrait-multimethod analysis. The age-10 BCS survey contains reports from both teacher and mother, with substantial overlap in items. Following Goodman et al. (2010), we use this to test the convergent and discriminant validity of our factor structure. If the underlying constructs are valid, mother and teacher should assess the child along the same dimensions: the correlation between, say, conduct measured from mother reports and conduct measured from teacher reports should be high (convergent validity), while the correlation between conduct measured from the mother and peer problems measured by the teacher should be lower (discriminant validity). For this exercise we take the raw sum of items from our main factorization, excluding items not included in the mother questionnaire from the teacher sum for comparability. Results are in Table B.4. The bottom left-hand block reports the relevant cross-correlations, with the diagonal element higher than the off-diagonal elements in most cases. The right-hand column reports formal tests for the marginal cases: the cross-correlations of attention with conduct, and of emotion with peer problems. As in Goodman et al. (2010), we find weaker discriminant validity between conduct and attention problems — a general issue with the SDQ conduct and hyperactivity scales rather than something specific to our analysis. Also consistent with their findings, the finer factorization performs better in terms of goodness of fit than the coarser externalising/internalising factorization.

Confirmatory factor analysis. Tables B.8 and B.9 report loadings from confirmatory factor analyses estimated on the main sample and the subsample with age-16 characteristics, respectively. Cross-loadings are set to zero in both cases. These CFA estimates provide the signal-to-noise ratios used in the measurement error correction described in Section A.6.

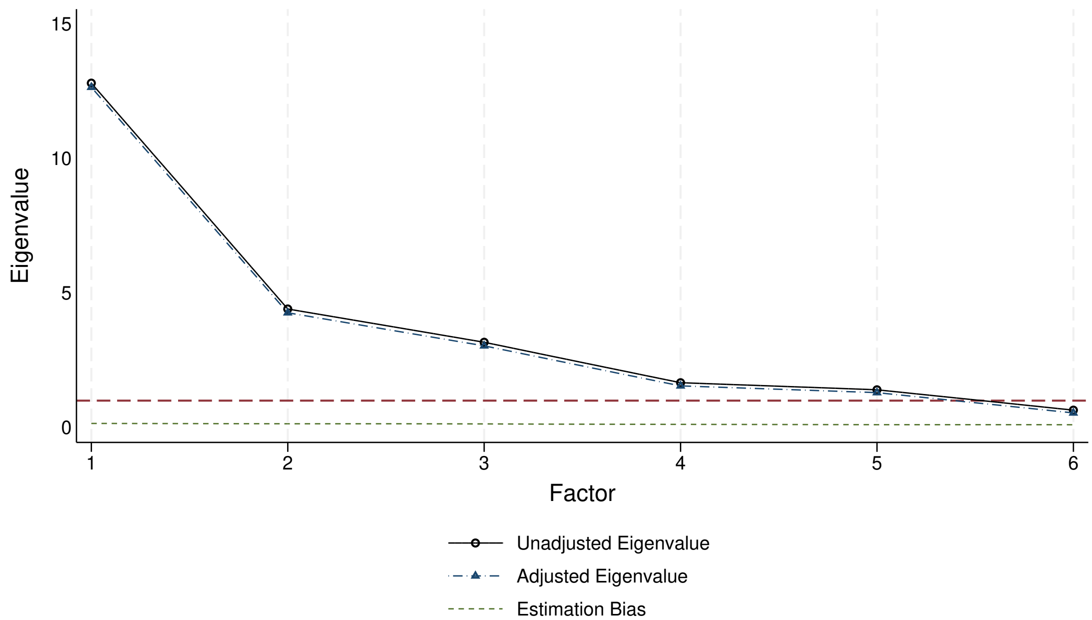
Correlations among factor scores. Table B.10 reports the correlation matrix of factor scores and other variables of interest. The correlation structure is discussed in the main text. One additional result worth noting: in a multivariate context (not shown), partialling out other factors, attention and conduct have no extra explanatory power for internalising behaviours beyond emotional and peer problems, while emotion and peer have no explanatory power for externalising after controlling for attention and conduct. This confirms that the four-factor structure cleanly nests the coarser externalising/internalising distinction.

Factorization of mother reports. Table B.5 reports an explicit factorization of mother reports, used in additional analyses in Section 6. Attention, conduct and emotional factors align closely with teacher reports; peer problems show weak internal consistency and no related factor emerges, consistent with the MTMM results above.

Locus of control and self-esteem. Table B.6 reports results from including 16 items from the Child Locus of Control (CARALOC) questionnaire and 12 items from the Lawrence Self-Esteem (LAWSEQ) questionnaire. Most locus-of-control items show weak loadings and are not retained.

The self-esteem items do generate an additional factor, but with Cronbach's alpha only just above 0.7, and the main text therefore retains the four socio-emotional plus one cognitive factor system as the preferred measurement system. Robustness results including self-esteem are in Appendix C.

Figure B.1: Eigenvalues from EFA of Age-10 Skills: Final Iteration



Notes: Data from BCS70. The figure shows unadjusted and adjusted eigenvalues from a factorization with retained items after several iterations of an exploratory factor analysis (see Table 1). Adjusted eigenvalues are obtained by using Horn's parallel factor analysis. This procedure involves performing factor analyses on both the actual data and samples generated randomly, extracting eigenvalues from each, and retaining only those factors whose eigenvalues exceed those expected by chance from random data.

Table B.1: Comparison with Labels from Strengths and Difficulties Questionnaire

SDQ Scales	Factorization	
	Attention	**Conduct**
Hyperactivity		
(i) Restless, overactive, cannot stay still for long		(i) Restless, over-active behaviour;
(ii) Constantly fidgeting or squirming		(ii) Hums or makes odd vocals; Rhythmic tapping in class
(iii) Easily distracted, concentration wanders	(iii) Easily distracted; Becomes bored during class; Child is daydreaming; Fails to pay attention in class; Fails to show perseverance; Shows lethargic behaviour	
(iv) Thinks things out before acting (reversed)		(iv) Excitable and impulsive
(v) Sees tasks through to the end, good attention span (reversed)	(v) Cannot complete tasks; Cannot concentrate on task; Forgetful on complex task; Fails to finish tasks	
Conduct problems		
(i) Often has temper tantrums or hot tempers		(i) Displays outbursts of temper; Changes mood quickly; Sullen or sulky
(ii) Generally obedient, usually does what adults request (reversed)		(ii) Destroys belongings; Easily frustrated; Complains about things
(iii) Often fights with other children or bullies them		(iii) Teases other children; Bullies other children; Quarrels with other kids; Interferes with others
(iv) Often lies or cheats		
(v) Steals from home, school or elsewhere		

Table B.1: Continued

SDQ Scales	Factorization	
	Emotion	Peer
Emotional problems		
(i) Often complains of headaches, stomach-aches or sickness	(i) Fussy	
(ii) Many worries, often seems worried	(ii) Worried	
(iii) Often unhappy, down-hearted or tearful		
(iv) Nervous or clingy in new situations, easily loses confidence	(iv) Afraid of new situations; Anxious; Behaves nervously	
(v) Many fears, easily scared		
Peer problems		
(i) Rather solitary, tends to play alone		(i) Rather solitary; Introvert
(ii) Has at least one good friend (reversed)		(ii) Child is not friendly
(iii) Generally liked by other children (reversed)		(iii) Child is not popular with peers
(iv) Picked on or bullied by other children		
(v) Gets on better with adults than with other children		
Prosocial		
(i) Considerate of other people's feelings		
(ii) Shares readily with other children (treats, toys, pencils etc.)		
(iii) Helpful if someone is hurt, upset or feeling ill		
(iv) Kind to younger children		
(v) Often volunteers to help others (parents, teachers, other children)		(v) Child is not cooperative

Notes: Data from BCS70. The table pairs items with similar definitions from 25 items in the Strengths & Difficulties Questionnaire (SDQ) and 42 items in the Child Development Questionnaire (BCS70). The 25 SDQ items comprise 5 scales (hyperactivity, conduct problems, emotional problems, peer problems, and prosocial) of 5 items each, which are then compared to items within the 4 socio-emotional skills derived in this paper.

Table B.2: EFA of Age-10 Skills for Restricted Item Set: Final Iteration

Items	**Externalising**	**Internalising**	**Cognition**	**Relationship to Main Factorization**
Bullies other children	**0.896**	-0.118	0.011	Conduct Problems
Quarrels with other kids	**0.809**	0.073	-0.013	Conduct Problems
Destroys belongings	**0.793**	0.042	-0.017	Conduct Problems
Restless or over-active behv.	**0.785**	0.060	0.000	Conduct Problems
Squirmy and fidgety	**0.760**	0.031	-0.123	Attention Problems*
Face or body twitches	**0.422**	0.267	0.035	Conduct Problems*
Worried	-0.030	**0.807**	-0.011	Emotion Problems
Unhappy or tearful	0.275	**0.671**	0.047	Emotion Problems*
Afraid of new situations	-0.205	**0.669**	-0.233	Emotion Problems
Rather solitary	-0.029	**0.625**	0.105	Peer Problems
Fussy	0.164	**0.518**	0.109	Emotion Problems
Child is not popular with peers	-0.231	**-0.513**	0.064	Peer Problems
Reading	-0.060	0.012	**0.839**	Cognition
Maths	-0.014	-0.042	**0.805**	Cognition
BAS words	0.027	0.019	**0.771**	Cognition
Pictorial (PLC)	0.009	0.028	**0.732**	Cognition
BAS simil	0.044	0.005	**0.728**	Cognition
BAS matrix	-0.086	0.029	**0.637**	Cognition
Spelling	-0.040	-0.045	**0.630**	Cognition
BAS digits	-0.003	-0.030	**0.452**	Cognition
Cronbach's alpha	0.788	0.719	0.857	

Notes: Data from BCS70. Items for the construction of externalising and internalising behaviours are obtained from the harmonized teacher questionnaire (Rutter Questionnaire). These are produced by Closer (https://discovery.closer.ac.uk/) and designed to be compatible with similar items from other surveys, such as the National Child Development Survey. The derived scores are given on a scale of 0-2. Table shows the final iteration of an exploratory factor analysis using a polychoric correlation matrix and oblique quartimin rotation. The last column shows the relationship between externalising and internalising behaviours with our 4 socio-emotional skills. Cronbach's alpha coefficients for the internal consistency of the set of retained items within each factor are also reported.
*Items with low loadings (<0.4) or high cross-loadings (>0.75) that were dropped after several iterations in our main factorization (see Table 1).

Table B.3: EFA of Age-10 Skills by Gender: Final Iteration

	Attention		Conduct		Emotion		Peer		Cognition	
	Males	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
Cannot complete tasks	**0.808**	**0.75**	-0.052	-0.053	-0.038	-0.061	0.05	0.029	-0.015	-0.024
Fails to finish tasks	**0.8**	**0.76**	-0.023	-0.041	-0.025	-0.018	0.05	0.026	0.009	-0.01
Easily distracted	**0.771**	**0.752**	0.167	0.165	0.022	0.052	-0.127	-0.124	-0.037	-0.044
Fails to show perseverance	**0.753**	**0.715**	0.031	0.009	-0.048	-0.051	0.024	0.03	-0.074	-0.049
Fails to pay attention in class	**0.722**	**0.754**	0.051	0.058	-0.081	-0.085	0.058	0.02	-0.098	-0.052
Daydreaming	**0.696**	**0.648**	-0.133	-0.12	0.123	0.124	0.104	0.102	0.041	0.018
Bored in class	**0.682**	**0.681**	0.17	0.163	-0.022	0	0.06	0.051	-0.025	0.004
Forgetful on complex task	**0.601**	**0.584**	-0.023	-0.006	0.199	0.186	-0.03	0.018	-0.204	-0.182
Cannot concentrate on task	**0.57**	**0.594**	0.019	-0.003	0.053	0.04	-0.014	0.031	-0.082	-0.062
Shows lethargic behaviour	**0.502**	**0.494**	-0.044	-0.018	0.135	0.139	0.273	0.268	-0.001	0.01
Displays outbursts of temper	-0.089	-0.07	**0.812**	**0.799**	0.037	0.034	0.084	0.052	-0.032	-0.015
Teases other children	-0.008	-0.021	**0.8**	**0.814**	-0.133	-0.124	0.021	0.037	-0.013	-0.051
Bullies other children	-0.04	-0.088	**0.769**	**0.824**	-0.151	-0.132	0.129	0.11	-0.071	-0.062
Quarrels with other kids	0.03	0.079	**0.767**	**0.74**	-0.039	0.004	0.166	0.125	-0.062	-0.033
Changes mood quickly	0.017	0.046	**0.71**	**0.708**	0.232	0.255	-0.005	-0.037	0.000	0.011
Interferes with others	0.278	0.214	**0.656**	**0.68**	-0.115	-0.099	-0.007	-0.019	-0.012	-0.025
Excitable and impulsive	0.067	0.088	**0.653**	**0.621**	0.214	0.219	-0.341	-0.328	0.023	0.015
Destroys belongings	0.063	0.069	**0.636**	**0.593**	-0.037	-0.104	0.134	0.178	-0.023	-0.014
Complains about things	0.046	0.118	**0.625**	**0.605**	0.087	0.115	0.05	-0.011	-0.008	0.011
Restless or over-active behv.	0.242	0.189	**0.599**	**0.6**	0.216	0.228	-0.26	-0.205	0.051	0.011
Sullen or sulky	0.023	0.065	**0.582**	**0.626**	0.094	0.09	0.247	0.224	-0.05	0.025
Easily frustrated	0.129	0.151	**0.544**	**0.545**	0.21	0.195	-0.073	-0.057	-0.041	0.019
Hums or makes odd vocals	0.215	0.152	**0.49**	**0.463**	0.056	-0.052	-0.077	0.043	0.019	-0.003
Rhythmic tapping in class	0.24	0.164	**0.47**	**0.441**	0.062	-0.015	-0.083	0.022	0.035	0.004
Worried	-0.016	0.004	0.078	0.095	**0.839**	**0.832**	0.004	-0.002	-0.022	-0.018
Anxious	-0.043	-0.054	-0.021	-0.005	**0.733**	**0.714**	0.151	0.207	-0.031	-0.05
Behaves nervously	0.094	0.074	-0.011	-0.043	**0.724**	**0.717**	0.072	0.088	-0.045	-0.074
Afraid of new situations	0.188	0.2	-0.155	-0.17	**0.534**	**0.563**	0.096	0.102	-0.172	-0.134
Fussy	-0.079	0.012	0.266	0.349	**0.526**	**0.515**	0.017	-0.098	0.081	0.062
Child is not friendly	0.135	0.118	0.152	0.147	0.097	0.059	**0.729**	**0.729**	-0.02	-0.045
Child is not popular with peers	0.159	0.148	0.196	0.213	0.067	0.025	**0.704**	**0.697**	-0.021	-0.048
Introvert	-0.024	-0.059	-0.34	-0.316	0.335	0.321	**0.584**	**0.594**	-0.01	-0.034
Child is not cooperative	0.131	0.144	0.35	0.334	-0.006	-0.023	**0.539**	**0.494**	-0.03	-0.001
Rather solitary	-0.002	-0.023	0.032	0.033	0.287	0.246	**0.532**	**0.568**	0.099	0.061
Reading	-0.071	-0.083	-0.015	0.016	0.001	0.01	0.006	0.006	**0.82**	**0.813**
BAS words	0.068	0.062	-0.003	-0.018	0.021	-0.003	-0.012	0.01	**0.806**	**0.793**
Pictorial (PLC)	0.114	0.084	-0.047	-0.049	0.006	-0.019	0.017	0.009	**0.788**	**0.752**
Maths	-0.123	-0.082	0.02	0.024	-0.035	-0.019	0.006	0.011	**0.766**	**0.77**
BAS simil	0.068	0.074	0.008	-0.026	-0.008	-0.014	0.024	0.003	**0.753**	**0.757**
BAS matrix	-0.054	0.012	-0.038	-0.04	0.007	0.021	0.036	-0.012	**0.637**	**0.647**
Spelling	-0.148	-0.225	0.061	0.067	-0.048	-0.009	-0.008	0.017	**0.597**	**0.536**
BAS digits	-0.057	-0.044	0.008	0.049	-0.041	-0.023	-0.005	0.005	**0.423**	**0.437**
Cronbach's alpha	0.926	0.916	0.931	0.929	0.847	0.848	0.832	0.815	0.861	0.854

Notes: Data from BCS70. The table compares factor loadings of retained items between males and females (3386 and 3566 observations, respectively) obtained from an exploratory factor analysis (EFA) and oblique quartimin rotation. Items are sorted in descending order of loadings within each construct for males. Cronbach's alpha coefficients for the internal consistency of the set of retained items within each factor are also reported.

Table B.4: MTMM Analysis of Age-10 Socio-Emotional Skills

		Teacher				Mother				Equal corr
		Attention	Conduct	Emotion	Peer	Attention	Conduct	Emotion	Peer	Prob > χ^2
Teacher	Attention	$\alpha = 0.79$								
	Conduct	0.57*	$\alpha = 0.91$							
	Emotion	0.33*	0.32*	$\alpha = 0.71$						
	Peer	0.30*	0.29*	0.42*	$\alpha = 0.62$					
Mother	Attention	**0.38***	**0.25***	0.10*	0.11*	$\alpha = 0.79$				0.00
	Conduct	**0.20***	**0.22***	0.03	0.09*	0.55*	$\alpha = 0.83$			0.21
	Emotion	-0.01	-0.05*	**0.15***	**0.07***	0.21*	0.33*	$\alpha = 0.59$		0.00
	Peer	0.05*	0.04	**0.08***	**0.18***	0.17*	0.31*	0.30*	$\alpha = 0.31$	0.00

Notes: Data from BCS70. Table shows a multitrait-multimethod analysis of age-10 social-emotional skills. The dataset comprises 19 aggregated measures that are close in definition reported by the pupils' teacher and mother (3 attention, 11 conduct, 3 emotion and 2 peer). Sample size for mother-teacher comparison is 6705. The diagonal shows Cronbach's alphas within each construct. The remaining cells contain Pairwise correlation coefficients (* p <.01, Bonferroni-adjusted significance level). The last column shows p-values associated with a chi-square test of equality of correlation between constructs in bold.

Table B.5: EFA of Age-10 Skills using Mother's Questionnaire: Final Iteration

Items	Attention	Conduct	Emotion	Cognition	Relationship to Main Factorization
Fails to finish things, short attention span	**0.661**	0.098	-0.077	-0.050	Attention
Very restless	**0.643**	-0.111	0.049	0.005	No close comparison
Cannot settle to do anything	**0.642**	0.115	-0.060	-0.057	No close comparison
Shows restless or overactive behaviour	**0.635**	0.008	0.161	0.083	Conduct
Inattentive, easily distracted	**0.622**	0.074	-0.023	-0.134	Attention
Squirmy or fidgety	**0.570**	-0.046	0.112	0.072	Attention*
Has difficulty concentrating on task	**0.564**	0.105	-0.008	-0.081	Attention
Is impulsive, excitable	**0.467**	-0.006	0.286	0.027	Conduct
Takes others' belongings	0.019	**0.635**	-0.040	0.004	No close comparison
Often tells lies	0.069	**0.609**	0.036	-0.026	No close comparison
Bullies other children	-0.034	**0.603**	0.095	0.001	Conduct
Destroys belongings	0.071	**0.562**	-0.014	-0.008	Conduct
Fights with other children	-0.001	**0.510**	0.090	-0.034	Conduct
Interferes with activity of other children	0.165	**0.461**	0.101	0.035	Conduct
Often disobedient	0.174	**0.425**	0.171	0.005	Conduct*
Worried	0.060	-0.201	**0.577**	-0.013	Emotion
Changes mood quickly and drastically	0.096	0.161	**0.553**	-0.032	Conduct
Appears miserable or distressed	-0.065	0.174	**0.536**	-0.020	Emotion*
Irritable	0.059	0.129	**0.531**	-0.009	No close comparison
Obsessed about unimportant things	0.089	-0.013	**0.521**	0.023	Emotion*
Requests must be met immediately	0.222	0.021	**0.519**	0.013	No close comparison
Is sullen or sulky	-0.078	0.200	**0.510**	-0.058	Conduct
Cries for little cause	-0.028	0.112	**0.505**	-0.064	Emotion*
Afraid of new things/situations	0.020	-0.161	**0.476**	-0.025	Emotion
Displays outbursts of temper	0.070	0.238	**0.473**	0.006	Conduct
Fussy over particular	-0.046	-0.094	**0.467**	-0.045	Emotion
Reading	-0.062	0.012	0.032	**0.841**	Cognition
Maths	-0.044	0.020	-0.009	**0.810**	Cognition
BAS words	0.069	-0.006	-0.048	**0.768**	Cognition
Pictorial (PLC)	0.077	-0.016	-0.042	**0.732**	Cognition
BAS simil	0.079	-0.004	-0.056	**0.727**	Cognition
BAS matrix	-0.055	-0.008	0.036	**0.639**	Cognition
Spelling	-0.120	0.012	0.072	**0.627**	Cognition
BAS digits	-0.021	-0.008	0.014	**0.452**	Cognition
Cronbach's alpha	0.848	0.801	0.857	0.8271	

Notes: Data from BCS70. Table shows the final iteration of an exploratory factor analysis using information from the mother's questionnaire (Rutter Questionnaire) on child's behaviour. The last column shows the relationship between socio-emotional domains obtained from this factorization with those obtained from the main factorization (See Table 1). Cronbach's alpha coefficients for the internal consistency of the set of retained items within each factor are also reported. * Items with low (<0.4) or high cross-loadings (>0.75) that were dropped after several iterations in the main factorization.

Table B.6: EFA of Age-10 Skills, Including Mindset: First Iteration

Items	Attention	Conduct	Emotion	Peer	Self-esteem	LOC	Cognition
Easily distracted	**0.787**	0.130	0.024	-0.113	-0.059	0.009	-0.046
Fails to finish tasks	**0.769**	-0.026	-0.020	0.063	0.005	-0.041	-0.003
Cannot complete tasks	**0.762**	-0.046	-0.049	0.074	-0.004	-0.019	-0.036
Fails to show perseverance	**0.717**	0.017	-0.064	0.060	-0.055	-0.004	-0.088
Fails to pay attention in class	**0.715**	0.058	-0.082	0.071	-0.023	-0.015	-0.100
Bored in class	**0.673**	0.165	-0.012	0.068	-0.013	-0.058	-0.010
Daydreaming	**0.654**	-0.121	0.115	0.122	-0.019	-0.047	0.027
Forgetful on complex task	**0.609**	-0.029	0.214	-0.003	0.005	-0.017	-0.208
Cannot concentrate on task	**0.579**	0.005	0.041	0.030	-0.022	-0.013	-0.082
Squirmy and fidgety	0.522	0.395	0.122	-0.188	-0.019	-0.041	0.051
Shows lethargic behaviour	**0.469**	0.004	0.150	0.283	0.042	-0.099	0.017
Confused with diffic. tasks	0.467	-0.080	0.387	-0.025	-0.012	-0.006	-0.265
Displays outbursts of temper	-0.057	**0.800**	0.036	0.063	-0.008	-0.015	-0.027
Bullies other children	-0.033	**0.791**	-0.128	0.094	0.035	-0.040	-0.068
Teases other children	0.029	**0.784**	-0.123	0.013	0.010	0.002	-0.038
Quarrels with other kids	0.048	**0.741**	-0.016	0.138	-0.100	0.054	-0.088
Changes mood quickly	0.024	**0.695**	0.246	-0.014	-0.047	0.019	-0.018
Destroys belongings	0.070	**0.635**	-0.027	0.127	0.070	-0.118	0.017
Interferes with others	0.287	**0.634**	-0.106	-0.024	-0.044	0.021	-0.030
Sullen or sulky	0.012	**0.624**	0.103	0.243	-0.023	-0.033	-0.028
Complains about things	0.061	**0.603**	0.113	0.016	-0.101	0.060	-0.040
Excitable and impulsive	0.126	**0.586**	0.203	-0.339	-0.058	0.062	0.008
Restless or over-active behv.	0.280	**0.560**	0.211	-0.249	-0.011	-0.018	0.062
Easily frustrated	0.158	**0.527**	0.204	-0.072	-0.025	-0.010	-0.014
Hums or makes odd vocals	0.237	**0.469**	0.027	-0.057	0.075	-0.156	0.093
Rhythmic tapping in class	0.242	**0.455**	0.054	-0.074	0.075	-0.167	0.100
Cannot negotiate child's behv.	0.272	0.320	-0.124	0.273	-0.003	-0.039	0.049
Face or body twitches	0.087	0.285	0.211	0.008	0.093	-0.140	0.085
Worried	0.007	0.065	**0.819**	0.002	-0.033	-0.002	-0.006
Behaves nervously	0.119	-0.032	**0.724**	0.066	0.037	-0.075	-0.013
Anxious	-0.032	-0.018	**0.696**	0.174	-0.025	-0.034	-0.013
Afraid of new situations	0.170	-0.173	**0.562**	0.099	-0.019	-0.018	-0.163
Fussy	-0.080	0.277	**0.545**	-0.050	-0.127	0.109	0.014
Obsessed with unimportant tasks	0.008	0.380	0.443	0.046	-0.038	0.003	-0.002
Cries for little cause	-0.045	0.322	0.409	0.081	-0.088	0.030	-0.045
Unhappy or tearful	-0.020	0.358	0.372	0.338	-0.048	-0.041	0.009
Child is not friendly	0.146	0.157	0.055	**0.704**	-0.106	-0.002	-0.033
Child is not popular with peers	0.164	0.209	0.025	**0.682**	-0.117	0.006	-0.044
Introvert	-0.046	-0.295	0.312	**0.592**	0.038	-0.097	0.016
Rather solitary	0.001	0.059	0.251	**0.535**	-0.008	-0.067	0.114
Child is not cooperative	0.149	0.356	-0.047	**0.530**	-0.062	0.004	-0.033
Child is not bold	0.009	-0.414	0.348	0.441	0.043	-0.067	-0.060
SEQ: sad, nobody to play with	0.032	0.076	-0.038	-0.023	**0.689**	-0.075	0.050
SEQ: feel lonely at school	-0.018	0.051	-0.056	-0.069	**0.670**	-0.022	-0.051
SEQ: children break friends	0.028	-0.037	0.050	-0.023	**0.627**	-0.004	0.087
SEQ: find friends, mine play with others	-0.008	0.009	0.006	-0.054	**0.620**	-0.035	0.073
SEQ: other children say nasty things	-0.029	-0.022	0.063	-0.023	**0.564**	0.084	-0.113
SEQ: people think that you tell lies	-0.037	-0.037	0.085	0.021	**0.472**	0.190	-0.096
SEQ: feel foolish when talking to teacher	0.044	0.074	-0.076	0.089	**0.429**	0.231	-0.054
LOC: blamed for things that aren't your fault	-0.086	-0.050	0.095	0.027	**0.425**	0.246	-0.111
SEQ: lots of things to change about yourself	-0.050	-0.003	0.040	0.030	**0.421**	0.133	-0.068

Table B.6: Continued

SEQ: feel foolish in front of other children	0.005	0.088	-0.054	0.066	0.382	0.175	0.021
SEQ: feel foolish when talking to parents	-0.018	0.005	0.030	0.074	0.361	0.158	0.028
SEQ: feel shy when talking to teachers	0.070	0.125	-0.103	0.070	0.352	0.118	0.040
LOC: feel sad when it's time to leave school	0.096	0.015	-0.045	-0.022	0.223	-0.194	0.200
LOC: studying for tests is a waste of time	-0.042	-0.049	0.052	-0.019	-0.010	**0.523**	-0.062
LOC: high marks just a matter of "luck"	0.033	-0.007	-0.011	-0.002	0.011	**0.507**	0.211
LOC: nice things happen is only good luck	0.099	-0.036	-0.002	0.001	0.034	**0.482**	0.206
LOC: tests are a lot of guess work	0.017	-0.017	-0.044	0.028	0.047	**0.480**	0.178
LOC: useless try in school, others are cleverer	-0.034	-0.013	-0.014	0.033	0.178	**0.471**	0.184
LOC: not worth trying, things never go well	0.040	-0.049	-0.001	0.042	0.176	**0.425**	0.106
LOC: get low marks, even when study hard	-0.141	0.066	0.010	-0.004	0.171	**0.401**	0.184
LOC: bad things, it's someone else's fault	-0.001	-0.024	0.048	-0.024	0.098	0.326	-0.073
SEQ: parents don't like to hear your ideas	-0.020	-0.024	0.048	-0.017	0.039	0.288	0.007
LOC: arguments, it's the other person's fault	0.014	0.000	0.021	-0.038	-0.018	0.190	-0.053
BAS words	0.078	-0.035	0.001	-0.002	-0.001	0.066	**0.771**
Reading	-0.149	-0.012	0.023	0.004	-0.009	0.046	**0.756**
Pictorial (PLC)	0.082	-0.066	-0.005	0.010	-0.005	0.028	**0.746**
Maths	-0.124	0.021	-0.015	0.000	0.054	0.016	**0.739**
BAS simil	0.064	-0.028	-0.010	0.008	-0.009	0.045	**0.726**
BAS matrix	-0.095	-0.042	0.035	0.001	-0.004	-0.013	**0.607**
Spelling	-0.250	0.059	-0.012	0.010	-0.008	0.090	**0.490**
BAS digits	-0.087	0.027	-0.024	0.011	0.004	0.059	0.384
LOC: people good no matter what	0.079	-0.003	-0.041	0.062	-0.182	-0.067	0.315
LOC: imp. make friends when angry	0.050	0.009	-0.029	0.004	0.123	0.113	0.280
LOC: surprised when you've done well	-0.050	0.034	-0.036	0.048	0.188	0.254	0.256
LOC: wishing make good things happen	0.064	-0.009	-0.032	0.040	0.097	-0.020	0.242
LOC: planning don't make things better	0.049	-0.038	0.010	-0.005	-0.029	0.064	0.096
Cronbach's alpha	0.925	0.931	0.847	0.823	0.711	0.667	0.857

Notes: Data from BCS70. The table reports factor loadings obtained from an exploratory factor analysis (EFA) of the main sample (6952 observations) and oblique quartimin rotation. Items in bold are retained after several iterations. We construct a dedicated measurement system adding self-esteem as an additional factor to our main specification, see equation (1). Table C.6 shows estimates including self-esteem on selected outcomes of interest. Cronbach's alpha coefficients for the internal consistency of the set of retained items within each factor are also reported.

Table B.7: EFAs of Family Background and Teen Socialization: First Iterations

Age10		Age16	
Items	**Family Backg.**	**Items**	**Teen Socialization**
Parental social status	**0.830**	**Go to a friend's house**	**0.785**
Parent in manager post	**0.717**	**Go out with friends**	**0.783**
Parental qualifications	**0.681**	**Have friends round to house**	**0.715**
Family income	**0.634**	**Spend time at friend's homes (school term)**	**0.686**
		Stay at home with friends (school term)	**0.633**
		Go out with friends do nothing special (school term)	**0.601**
		Hang about the street	**0.559**
		Go with friends to cinema, disco etc. (school term)	**0.466**
		Go to a youth club/organization	0.303
		Number of friends	0.258
		Go out with brothers/sisters	0.107
		Go to a meeting/political club	-0.018
Eigenvalues	2.068		3.673
Cronbach's alpha	0.712		0.790

Notes: Data from BCS70. The table reports factor loadings obtained from an exploratory factor analysis (EFA) of the main sample for family background (6952 observations) and age-16 sample for teen socialization (3377 observations) using a polychoric correlation matrix and oblique quartimin rotation. Items in bold are retained after several iterations and used in our dedicated measurement system. Eigenvalues and Cronbach's alpha coefficients for the internal consistency of the set of retained items within each factor are also reported.

Table B.8: Loadings from Confirmatory Factor Analysis: Main Sample

Items	Loadings	Items	Loadings
Attention Problems		**Emotion Problems**	
Easily distracted	0.85	Worried	0.87
Bored in class	0.80	Anxious	0.80
Fails to pay attention in class	0.79	Behaves nervously	0.80
Fails to show perseverance	0.78	Afraid of new situations	0.67
Fails to finish tasks	0.77	Fussy	0.51
Cannot complete tasks	0.76		
Forgetful on complex task	0.75	**Peer Problems**	
Daydreaming	0.66	Child is not friendly	0.91
Cannot concentrate on task	0.65	Child is not popular with peers	0.91
Shows lethargic behaviour	0.61	Child is not cooperative	0.68
		Rather solitary	0.56
		Introvert	0.45
Conduct Problems		**Cognition**	
Quarrels with other kids	0.80	Reading	0.86
Interferes with others	0.79	Maths	0.83
Displays outbursts of temper	0.78	BAS words	0.74
Teases other children	0.77	Pictorial (PLC)	0.72
Changes mood quickly	0.77	BAS simil	0.70
Bullies other children	0.75	BAS matrix	0.66
Restless or over-active behv.	0.73	Spelling	0.66
Sullen or sulky	0.68	BAS digits	0.46
Complains about things	0.68		
Easily frustrated	0.67	**Family SES**	
Destroys belongings	0.67	Parental social status	0.83
Excitable and impulsive	0.66	Parental qualifications	0.69
Hums or makes odd vocals	0.58	Family income	0.59
Rhythmic tapping in class	0.58	Parent in manager post	0.50
Tucker-Lewis index (TLI)		0.988	
Comparative fit index (CFI)		0.804	
Std root-mean-sq-residual (RMSR)		0.068	
Root mean sq error of approx (RMSEA)		0.101	

Notes: Data from BCS70. The table shows estimated factor loadings from a confirmatory factor analysis (CFA) for the measurement system described in equation (1) using a sample of 6952 observations. Incremental and absolute measures of goodness of fit are also reported: TLI (Tucker & Lewis, 1973) and CFI (Bentler, 1990) have a scale from zero to one, with values close to one indicating a good fit. RMSR (Jöreskog & Sörbom, 1986) and RMSEA (Steiger, 1990) also range from zero to one, with values close to zero indicating a good fit.

Table B.9: Loadings from CFA: Sample with Age-16 Characteristics

Items	Loadings	Items	Loadings
Teen Socialization		**Emotion Problems**	
Go to a friend's house	0.75	Worried	0.87
Have friends round to my house	0.70	Behaves nervously	0.81
Go out with friends	0.69	Anxious	0.79
Spend time at friends' homes (school term)	0.64	Afraid of new situations	0.68
Stay at home with friends (school term)	0.59	Fussy	0.53
Go out with friends do nothing special (school term)	0.47		
Go with friends to cinema, disco etc. (school term)	0.41		
Hang about the streets	0.37	**Peer Problems**	
		Child is not friendly	0.91
Attention Problems		Child is not popular with peers	0.91
Easily distracted	0.85	Child is not cooperative	0.68
Bored in class	0.79	Rather solitary	0.58
Fails to pay attention in class	0.77	Introvert	0.47
Forgetful on complex task	0.76		
Fails to show perseverance	0.76		
Fails to finish tasks	0.75	**Cognition**	
Cannot complete tasks	0.73	Reading	0.85
Daydreaming	0.67	Maths	0.82
Cannot concentrate on task	0.63	BAS words	0.74
Shows lethargic behaviour	0.62	Pictorial (PLC)	0.71
		BAS simil	0.70
Conduct Problems		BAS matrix	0.65
Quarrels with other kids	0.79	Spelling	0.63
Changes mood quickly	0.78	BAS digits	0.44
Displays outbursts of temper	0.77		
Interferes with others	0.77		
Teases other children	0.75	**Family SES**	
Restless or over-active behv.	0.74	Parental social status	0.83
Bullies other children	0.74	Parental qualifications	0.71
Complains about things	0.69	Family income	0.61
Sullen or sulky	0.69	Parent in manager post	0.49
Destroys belongings	0.67		
Easily frustrated	0.67		
Excitable and impulsive	0.66		
Hums or makes odd vocals	0.55		
Rhythmic tapping in class	0.57		
Tucker-Lewis index (TLI)		0.769	
Comparative fit index (CFI)		0.781	
Std root-mean-sq-residual (RMSR)		0.065	
Root mean sq error of approx (RMSEA)		0.063	

Notes: Data from BCS70. The table shows estimated factor loadings from a Confirmatory Factor Analysis (CFA) for the measurement system described in equation (1) adding teen socialization as an additional latent construct and using a sample of 3377 observations. Incremental and absolute measures of fit are also reported: TLI (Tucker & Lewis, 1973) and CFI (Bentler, 1990) have a scale from zero to one, with values close to one indicating a good fit. RMSR (Jöreskog & Sörbom, 1986) and RMSEA (Steiger, 1990) also range from zero to one, with values close to zero indicating a good fit.

Table B.10: Correlation Matrix: Age-10 Skills and Other Background Characteristics

	Atte	**Cond**	**Emot**	Peer	Exte	Inte	SEst	Cogn	FSES	Soci	YSch
Attention	1.00										
Conduct	0.59	1.00									
Emotion	0.39	0.29	1.00								
Peer	0.46	0.38	0.45	1.00							
Externalising	**0.60**	**0.89**	0.27	0.35	1.00						
Internalising	0.45	0.41	**0.84**	**0.69**	0.38	1.00					
Self-Esteem	-0.19	-0.16	-0.17	-0.23	-0.15	-0.21	1.00				
Cognition	-0.48	-0.18	-0.25	-0.25	-0.21	-0.25	0.24	1.00			
Family SES	-0.19	-0.11	-0.10	-0.10	-0.09	-0.10	0.14	0.39	1.00		
Socialization	0.02	0.02	-0.06	-0.09	0.02	-0.08	0.01	-0.06	-0.06	1.00	
Years School	-0.32	-0.15	-0.13	-0.14	-0.15	-0.13	0.16	0.50	0.42	-0.13	1.00

Notes: Data from BCS70. Factor scores are obtained from a confirmatory factor analysis (CFA) of a dedicated measurement system, see equation (1). The table shows the correlation matrix among factor scores reported in this study and years of schooling. Externalising and internalising behaviours are factor scores obtained from a dedicated measurement system using items from the harmonized teacher questionnaire (see Table B.2).

C Additional Results Quoted in the Main Paper

Table C.1: Earnings Determinants, Controlling for Occupational Sorting

	All		Males	Females	q-values
	[1]	[2]	[3]	[4]	[3]–[4]
Attention	-0.027^{***}	-0.015^{*}	-0.017^{*}	-0.013	1
	[0.009]	[0.008]	[0.009]	[0.011]	
Conduct	0.037^{***}	0.024^{***}	0.024^{***}	0.022^{***}	1
	[0.006]	[0.005]	[0.007]	[0.008]	
Emotion	-0.029^{***}	-0.021^{***}	-0.018^{**}	-0.023^{***}	1
	[0.006]	[0.005]	[0.007]	[0.008]	
Peer	-0.014^{**}	-0.007	-0.022^{***}	0.008	0.030
	[0.007]	[0.006]	[0.008]	[0.008]	
Cognition	0.060^{***}	0.026^{***}	0.017^{**}	0.032^{***}	0.707
	[0.007]	[0.006]	[0.008]	[0.009]	
Family SES	0.046^{***}	0.030^{***}	0.048^{***}	0.008	0.004
	[0.007]	[0.006]	[0.008]	[0.008]	
Yrs School	0.049^{***}	0.027^{***}	0.016^{***}	0.041^{***}	0.001
	[0.003]	[0.003]	[0.003]	[0.005]	
Occupation dummies		X	X	X	
Backg. controls	X	X	X	X	
Mean of Dep. Var.	7.17	7.17	7.44	6.92	
N individuals	6,952	6,952	3,386	3,566	
N individual-years	23,451	23,451	11,404	12,047	

Notes: Data from BCS70. Each column reports measurement-error-corrected estimates from a regression of log real monthly earnings as the dependent variable. See notes to Table 2. Specifications [2], [3] and [4] adds UK-SOC2010 (3-digits) occupational dummies as controls. All specifications include controls for: number of siblings, dummies for first child, no dad at birth, teenage mother, year and region fixed effects. Specifications [1] and [2] also include controls for gender and gender-by-year. Standard errors in brackets are estimated from 250 bootstrap replications clustered at the individual level: ***p < .01, **p < .05, *p < .10. FDR q-values report gender differences adjusted for multiple hypothesis testing across years of schooling, earnings, wages, working hours and employment.

Table C.2: Gender Differences in Schooling and Earnings

	Schooling			Earnings		
	Males [1]	**Females** [2]	**q-values** [1]–[2]	**Males** [3]	**Females** [4]	**q-values** [3]–[4]
Attention	-0.307^{***}	-0.114^{**}	0.055	-0.033^{***}	-0.043^{***}	1
	[0.066]	[0.054]		[0.011]	[0.015]	
Conduct	-0.025	-0.052	1	0.027^{***}	0.042^{***}	0.707
	[0.051]	[0.045]		[0.008]	[0.011]	
Emotion	0.078	0.034	1	-0.024^{***}	-0.029^{***}	1
	[0.051]	[0.041]		[0.008]	[0.010]	
Peer	0.022	0.028	1	-0.028^{***}	0.007	0.041
	[0.050]	[0.040]		[0.009]	[0.011]	
Cognition	0.724^{***}	0.730^{***}	1	0.061^{***}	0.126^{***}	0.001
	[0.062]	[0.045]		[0.010]	[0.012]	
Family SES	0.595^{***}	0.624^{***}	1	0.075^{***}	0.073^{***}	1
	[0.055]	[0.044]		[0.009]	[0.010]	
Backg. controls	X	X		X	X	
Mean of Dep. Var.	12.29	12.23		7.44	6.92	
N individuals	3,386	3,566		3,386	3,566	
N individual-years				11,404	12,047	

Notes: Data from BCS70. Columns [1]–[2] report gender-specific regressions of years of schooling; columns [3]–[4] report regressions of log monthly earnings. Outcomes are regressed on standardised socio-emotional skills, cognition, and family socio-economic status (SES), constructed using our dedicated measurement system (see equation 1). All specifications include controls for number of siblings, dummies for first child, no father at birth, teenage mother, and region fixed effects. Earnings models additionally include year fixed effects. Standard errors in square brackets are estimated from 250 bootstrap replications clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. FDR q-values report gender differences adjusted for multiple hypothesis testing across years of schooling, earnings, wages, working hours and employment.

Table C.3: Gender Differences in Wages, Hours, and Employment

	Wages			Hours			Employment		
	Males	**Females**	**q-val**	**Males**	**Females**	**q-val**	**Males**	**Females**	**q-val**
	[1]	[2]	[1]–[2]	[3]	[4]	[3]–[4]	[5]	[6]	[5]–[6]
Attention	-0.036^{***}	-0.036^{***}	1	0.003	-0.007	0.806	-0.002	-0.030^{***}	0.030
	[0.010]	[0.010]		[0.004]	[0.010]		[0.006]	[0.008]	
Conduct	0.019^{**}	0.031^{***}	0.741	0.008^{**}	0.012^{*}	1	-0.017^{***}	0.012^{**}	0.002
	[0.008]	[0.008]		[0.003]	[0.007]		[0.005]	[0.006]	
Emotion	-0.017^{**}	-0.005	0.707	-0.007^{**}	-0.024^{***}	0.055	-0.003	-0.002	1
	[0.007]	[0.007]		[0.003]	[0.007]		[0.004]	[0.006]	
Peer	-0.020^{***}	-0.011	0.806	-0.007^{**}	0.018^{***}	0.004	-0.010^{**}	-0.003	0.806
	[0.008]	[0.007]		[0.003]	[0.007]		[0.004]	[0.006]	
Cognition	0.067^{***}	0.086^{***}	0.398	-0.006^{**}	0.041^{***}	0.001	0.024^{***}	0.023^{***}	1
	[0.010]	[0.008]		[0.003]	[0.007]		[0.006]	[0.007]	
Family SES	0.070^{***}	0.066^{***}	1	0.006^{*}	0.007	1	0.003	0.021^{***}	0.037
	[0.009]	[0.007]		[0.003]	[0.006]		[0.005]	[0.006]	
Backg. controls	X	X		X	X		X	X	
Mean Dep. Var.	2.26	2.06		5.18	4.86		0.91	0.79	
N individuals	3,386	3,566		3,386	3,566		4,017	4,145	
N ind.-years	11,404	12,047		11,404	12,047		16,278	18,335	

Notes: Data from BCS70. Each column reports measurement-error-corrected estimates from gender-specific regressions of log hourly wages (columns [1]–[2]), log monthly working hours (columns [3]–[4]), and whether an individual is employed (columns [5]–[6]) on standardised socio-emotional skills, cognition, and family socio-economic status (SES), constructed using our dedicated measurement system (see equation 1). The employment outcome is estimated via a linear probability model (LPM) using all survey participants with valid observations across the 1996–2016 waves. All specifications include controls for number of siblings, dummies for first child, no father at birth, teenage mother, and year and region fixed effects. Standard errors in square brackets are estimated from 250 bootstrap replications clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. FDR q-values report gender differences adjusted for multiple hypothesis testing across years of schooling, earnings, wages, working hours and employment.

Table C.4: Attachment to the BCS Survey

	All	Males	Females
	[1]	[2]	[3]
Attention	0.001	-0.061	0.081
	[0.039]	[0.051]	[0.058]
Conduct	-0.083^{**}	-0.088^{**}	-0.092^{*}
	[0.032]	[0.043]	[0.048]
Emotion	0.072^{**}	0.016	0.148^{***}
	[0.030]	[0.040]	[0.045]
Peer	-0.036	0.014	-0.085^{*}
	[0.032]	[0.043]	[0.046]
Cognition	0.285^{***}	0.230^{***}	0.364^{***}
	[0.033]	[0.044]	[0.050]
Family SES	0.076^{***}	0.070^{*}	0.081^{*}
	[0.028]	[0.037]	[0.043]
Backg. controls	X	X	X
Mean of Dep. Var.	0.70	0.64	0.76
N individuals	6,952	3,386	3,566

Notes: Data from BCS70. The table reports estimates from logit regressions of full survey participation, defined as responding in all waves from 1996 to 2016, regardless of labour market status on age-10 socio-emotional skills, cognition, and family socio-economic status (SES), constructed using our dedicated measurement system (see equation 1). The sample comprises the 6,952 individuals used in the main analysis. Columns [2] and [3] report gender-specific estimates. All specifications include controls for number of siblings, dummies for first child, no father at birth, teenage mother, and region fixed effects. Column [1] additionally includes gender and gender-by-year controls. Standard errors in square brackets: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table C.5: Main Regressions, with Inverse Probability Weighting

	Schooling		Earnings		Wages		Hours	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Attention	-0.221^{***}	-0.216^{***}	-0.037^{***}	-0.035^{***}	-0.037^{***}	-0.035^{***}	-0.001	-0.001
	[0.045]	[0.045]	[0.009]	[0.009]	[0.007]	[0.007]	[0.005]	[0.005]
Conduct	-0.037	-0.037	0.036^{***}	0.033^{***}	0.025^{***}	0.023^{***}	0.010^{***}	0.010^{***}
	[0.036]	[0.036]	[0.007]	[0.007]	[0.005]	[0.005]	[0.004]	[0.003]
Emotion	0.058^{*}	0.056^{*}	-0.028^{***}	-0.027^{***}	-0.011^{**}	-0.011^{**}	-0.017^{***}	-0.016^{***}
	[0.033]	[0.033]	[0.007]	[0.006]	[0.005]	[0.005]	[0.003]	[0.003]
Peer	0.025	0.020	-0.011	-0.014^{*}	-0.016^{***}	-0.017^{***}	0.005	0.004
	[0.032]	[0.032]	[0.007]	[0.007]	[0.006]	[0.006]	[0.004]	[0.004]
Cognition	0.726^{***}	0.677^{***}	0.093^{***}	0.090^{***}	0.077^{***}	0.075^{***}	0.016^{***}	0.015^{***}
	[0.036]	[0.038]	[0.007]	[0.007]	[0.006]	[0.006]	[0.004]	[0.004]
Family SES	0.610^{***}	0.592^{***}	0.075^{***}	0.073^{***}	0.068^{***}	0.067^{***}	0.007^{*}	0.006^{*}
	[0.032]	[0.034]	[0.007]	[0.007]	[0.005]	[0.005]	[0.004]	[0.004]
Backg. controls	X	X	X	X	X	X	X	X
Inv. prob. weight.		X		X		X		X
Mean of Dep. Var.	12.26	12.26	7.17	7.17	2.16	2.16	5.02	5.02
N individuals	6,952	6,952	6,952	6,952	6,952	6,952	6,952	6,952
N individual-years			23,451	23,451	23,451	23,451	23,451	23,451

Notes: Data from BCS70. The table compares measurement-error-corrected estimates from regressions of years of schooling, log monthly earnings, log hourly wages, and log monthly working hours on standardised socio-emotional skills, cognition, and family socio-economic status (SES), with and without inverse probability weighting (IPW). Odd-numbered columns report unweighted estimates; even-numbered columns re-estimate the same specifications using IPW. Weights are estimated from a logit regression of full survey participation on age-10 skills and socio-demographic characteristics (see Table C.4). All specifications include controls for gender, gender-by-year, number of siblings, dummies for first child, no father at birth, teenage mother, and year and region fixed effects. Standard errors in square brackets are estimated from 250 bootstrap replications clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table C.6: Main Regressions, Controlling for Self-Esteem

	Schooling		Earnings		Wages		Hours	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Self-Esteem		0.060^{**}		0.010		0.018^{***}		-0.007^{*}
		[0.029]		[0.007]		[0.005]		[0.004]
Attention	-0.221^{***}	-0.219^{***}	-0.037^{***}	-0.037^{***}	-0.037^{***}	-0.036^{***}	-0.001	-0.001
	[0.045]	[0.045]	[0.009]	[0.009]	[0.007]	[0.007]	[0.005]	[0.005]
Conduct	-0.037	-0.032	0.036^{***}	0.036^{***}	0.025^{***}	0.027^{***}	0.010^{***}	0.010^{***}
	[0.036]	[0.036]	[0.007]	[0.007]	[0.005]	[0.005]	[0.004]	[0.004]
Emotion	0.058^{*}	0.058^{*}	-0.028^{***}	-0.028^{***}	-0.011^{**}	-0.011^{**}	-0.017^{***}	-0.017^{***}
	[0.033]	[0.033]	[0.007]	[0.007]	[0.005]	[0.005]	[0.003]	[0.003]
Peer	0.025	0.036	-0.011	-0.010	-0.016^{***}	-0.013^{**}	0.005	0.004
	[0.032]	[0.032]	[0.007]	[0.008]	[0.006]	[0.006]	[0.004]	[0.004]
Cognition	0.726^{***}	0.716^{***}	0.093^{***}	0.091^{***}	0.077^{***}	0.074^{***}	0.016^{***}	0.018^{***}
	[0.036]	[0.037]	[0.007]	[0.007]	[0.006]	[0.006]	[0.004]	[0.004]
Family SES	0.610^{***}	0.607^{***}	0.075^{***}	0.075^{***}	0.068^{***}	0.067^{***}	0.007^{*}	0.008^{*}
	[0.032]	[0.032]	[0.007]	[0.007]	[0.005]	[0.005]	[0.004]	[0.004]
Backg. controls	X	X	X	X	X	X	X	X
Mean of Dep. Var.	12.26	12.26	7.17	7.17	2.16	2.16	5.02	5.02
N individuals	6,952	6,952	6,952	6,952	6,952	6,952	6,952	6,952
N individual-years			23,451	23,451	23,451	23,451	23,451	23,451

Notes: Data from BCS70. Odd-numbered columns reproduce the baseline measurement-error-corrected estimates from regressions of years of schooling, log monthly earnings, log hourly wages, and log monthly working hours on standardised socio-emotional skills, cognition, and family socio-economic status (SES). Even-numbered columns augment each specification with a self-esteem factor, constructed alongside the main skills using our dedicated measurement system (see equation 1). All specifications include controls for gender, gender-by-year, number of siblings, dummies for first child, no father at birth, teenage mother, and year and region fixed effects. Standard errors in square brackets are estimated from 250 bootstrap replications clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table C.7: Earnings Determinants, Including Skill Interactions

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Attention	-0.037^{***}	-0.041^{***}	-0.040^{***}	-0.041^{***}	-0.039^{***}	-0.039^{***}	-0.040^{***}
	[0.009]	[0.007]	[0.007]	[0.007]	[0.007]	[0.007]	[0.007]
Conduct	0.036^{***}	0.043^{***}	0.031^{***}	0.038^{***}	0.031^{***}	0.031^{***}	0.031^{***}
	[0.007]	[0.008]	[0.006]	[0.007]	[0.006]	[0.006]	[0.006]
Emotion	-0.028^{***}	-0.027^{***}	-0.025^{***}	-0.026^{***}	-0.025^{***}	-0.025^{***}	-0.025^{***}
	[0.007]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]
Peer	-0.011	-0.014^{**}	-0.015^{**}	-0.013^{**}	-0.015^{**}	-0.014^{**}	-0.015^{**}
	[0.007]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]
Cognition	0.093^{***}	0.099^{***}	0.100^{***}	0.100^{***}	0.100^{***}	0.100^{***}	0.100^{***}
	[0.007]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]
Conduct $\times$ Attention		-0.016^{***}					
		[0.005]					
Conduct $\times$ Emotion			0.000				
			[0.005]				
Conduct $\times$ Peer				-0.010^{**}			
				[0.004]			
Attention $\times$ Emotion					-0.003		
					[0.005]		
Attention $\times$ Peer						-0.004	
						[0.005]	
Emotion $\times$ Peer							-0.003
							[0.005]
Family SES	X	X	X	X	X	X	X
Backg. controls	X	X	X	X	X	X	X
Mean of Dep. Var.	7.17	7.17	7.17	7.17	7.17	7.17	7.17
N individuals	6,952	6,952	6,952	6,952	6,952	6,952	6,952
N individual-years	23,451	23,451	23,451	23,451	23,451	23,451	23,451

Notes: Data from BCS70. Each column reports estimates from a regression of log monthly earnings on standardised socio-emotional skills, cognition, and family socio-economic status (SES), constructed using our dedicated measurement system (see equation 1). One interaction term between socio-emotional skills is introduced at a time. All specifications include controls for gender, gender-by-year, number of siblings, dummies for first child, no father at birth, teenage mother, and year and region fixed effects. Standard errors in square brackets are clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table C.8: Determinants of Earnings, Including Cognition Interactions

	Earnings			
	[1]	[2]	[3]	[4]
Attention	-0.043^{***}	-0.039^{***}	-0.040^{***}	-0.040^{***}
	[0.007]	[0.007]	[0.007]	[0.007]
Conduct	0.031^{***}	0.030^{***}	0.031^{***}	0.031^{***}
	[0.006]	[0.006]	[0.006]	[0.006]
Emotion	-0.024^{***}	-0.025^{***}	-0.025^{***}	-0.025^{***}
	[0.006]	[0.006]	[0.006]	[0.006]
Peer	-0.015^{**}	-0.015^{**}	-0.015^{**}	-0.015^{**}
	[0.006]	[0.006]	[0.006]	[0.006]
Cognition	0.101^{***}	0.100^{***}	0.100^{***}	0.100^{***}
	[0.006]	[0.006]	[0.006]	[0.006]
Cognition × Attention	-0.011^{**}			
	[0.005]			
Cognition × Conduct		-0.005		
		[0.005]		
Cognition × Emotion			0.001	
			[0.005]	
Cognition × Peer				-0.001
				[0.005]
Family SES	X	X	X	X
Backg. controls	X	X	X	X
Mean of Dep. Var.	7.17	7.17	7.17	7.17
N individuals	6,952	6,952	6,952	6,952
N individual-years	23,451	23,451	23,451	23,451

Notes: Data from BCS70. Each column reports estimates from a regression of log monthly earnings on standardised socio-emotional skills, cognition, and family socio-economic status (SES), constructed using our dedicated measurement system (see equation 1). One interaction term between cognition and a socio-emotional trait is introduced at a time. All specifications include controls for gender, gender-by-year, number of siblings, dummies for first child, no father at birth, teenage mother, and year and region fixed effects. Standard errors in square brackets are clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table C.9: The Role of Teacher vs Mother Assessments of Skills

	Schooling			Earnings		
	[1]	[2]	[3]	[4]	[5]	[6]
Mother assessed:						
Attention		-0.142^{***}	-0.080^{*}		-0.013	-0.007
		[0.042]	[0.043]		[0.009]	[0.007]
Conduct		-0.069	-0.044		0.018^{**}	0.004
		[0.045]	[0.046]		[0.009]	[0.006]
Emotion		0.114^{**}	0.064		-0.003	-0.001
		[0.047]	[0.049]		[0.009]	[0.007]
Teacher assessed:						
Attention	-0.221^{***}		-0.211^{***}	-0.037^{***}		-0.040^{***}
	[0.045]		[0.046]	[0.009]		[0.008]
Conduct	-0.037		-0.026	0.036^{***}		0.033^{***}
	[0.036]		[0.036]	[0.007]		[0.006]
Emotion	0.058^{*}		0.048	-0.028^{***}		-0.023^{***}
	[0.033]		[0.035]	[0.007]		[0.006]
Peer	0.025		0.029	-0.011		-0.016^{**}
	[0.032]		[0.033]	[0.007]		[0.006]
Other:						
Cognition	0.726^{***}	0.804^{***}	0.709^{***}	0.093^{***}	0.115^{***}	0.099^{***}
	[0.036]	[0.029]	[0.036]	[0.007]	[0.006]	[0.006]
Family SES	0.610^{***}	0.603^{***}	0.619^{***}	0.075^{***}	0.075^{***}	0.068^{***}
	[0.032]	[0.034]	[0.033]	[0.007]	[0.007]	[0.006]
Backg. controls	X	X	X	X	X	X
Mean of Dep. Var.	12.26	12.26	12.26	7.17	7.17	7.17
N individuals	6,952	6,692	6,692	6,952	6,692	6,692
N individual-years				23,451	22,642	22,642

Notes: Data from BCS70. The table compares measurement-error-corrected estimates from regressions of years of schooling and log monthly earnings on our four-factor teacher-assessed model (columns [1] and [4]) with a three-factor model of socio-emotional skills assessed by mothers (columns [2] and [5]; see Table B.5), and a combined specification including both (columns [3] and [6]). All socio-emotional skills, cognition, and family socio-economic status (SES) are standardised measures constructed using our dedicated measurement system (see equation 1). All specifications include controls for gender, gender-by-year, number of siblings, dummies for first child, no father at birth, teenage mother, and year and region fixed effects. Standard errors in square brackets are estimated from 250 bootstrap replications clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table C.10: Determinants of Schooling and Earnings: Inclusion of Extra Controls

	Schooling				Earnings			
Attention	-0.245^{***}	-0.246^{***}	-0.248^{***}	-0.260^{***}	-0.036^{***}	-0.035^{***}	-0.036^{***}	-0.037^{***}
	[0.059]	[0.059]	[0.059]	[0.060]	[0.010]	[0.010]	[0.010]	[0.010]
Conduct	-0.018	-0.014	-0.008	-0.005	0.040^{***}	0.040^{***}	0.040^{***}	0.042^{***}
	[0.044]	[0.043]	[0.043]	[0.043]	[0.009]	[0.009]	[0.009]	[0.009]
Emotion	0.064	0.057	0.054	0.052	-0.015^{*}	-0.016^{*}	-0.016^{*}	-0.016^{*}
	[0.043]	[0.042]	[0.043]	[0.043]	[0.008]	[0.008]	[0.008]	[0.008]
Peer	0.058	0.058	0.055	0.053	-0.007	-0.007	-0.007	-0.008
	[0.042]	[0.041]	[0.041]	[0.041]	[0.009]	[0.009]	[0.009]	[0.009]
Cognition	0.748^{***}	0.726^{***}	0.721^{***}	0.706^{***}	0.102^{***}	0.099^{***}	0.099^{***}	0.097^{***}
	[0.048]	[0.048]	[0.048]	[0.049]	[0.009]	[0.009]	[0.009]	[0.009]
Family SES	0.585^{***}	0.527^{***}	0.512^{***}	0.481^{***}	0.069^{***}	0.061^{***}	0.062^{***}	0.061^{***}
	[0.045]	[0.051]	[0.053]	[0.055]	[0.009]	[0.010]	[0.010]	[0.010]
Backg. Controls	X	X	X	X	X	X	X	X
Neighbourhood SES & Housing cond.		X	X	X		X	X	X
Parental warmth & Early-life health			X	X			X	X
School SES & Char. + Discipline				X				X
Mean of Dep. Var.	12.28	12.28	12.28	12.28	7.17	7.17	7.17	7.17
R squared	0.283	0.289	0.292	0.297	0.343	0.344	0.344	0.346
N individual	4,121	4,121	4,121	4,121	4,121	4,121	4,121	4,121
N individual-years					14,058	14,058	14,058	14,058

Notes: Data from BCS70. Each column reports measurement-error-corrected regression estimates from regressions of years of schooling and log monthly earnings on standardized socio-emotional and cognitive skills obtained from our dedicated measurement system (see equation 1). Each column sequentially introduces sets of control variables. **Neighbourhood SES & housing conditions** include variables such as the social rating of neighbourhood, tenure of accommodation, and the persons-per-room ratio, all measured at age 5. **Parental warmth** is a factor score based on the Maternal Self-Completion Questionnaire at age 5, capturing non-authoritarian child-rearing practices and broader worldviews. **Early-life health** includes variables such as birth-weight (in ounces), gestation (premature or not), and whether the child was breastfed, all recorded at age 5. **School characteristics** include school type, class size, gender composition, teacher responsibility for behavioural problems, presence of a teaching assistant, assistant's working hours, and availability of special education support, all reported at age 10. **School SES** captures the socio-economic composition of the school, based on parental academic level, occupation, and neighbourhood characteristics as reported by the teacher at age 10. Finally, **Discipline** is a standardized score based on 13 practices, each measured on a 4-point scale from "often" to "never" (e.g., suspension, corporal punishment, physical exercise, extra homework), also reported at age 10. All specifications control for gender, gender-by-year, number of siblings, dummies for first child, no dad at birth, teenage mother, year and region fixed effects. Standard errors in square brackets are estimated from 250 bootstrap replications, clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table C.11: Determinants of Teen Characteristics by Gender

	Males				Females			
	Sport Comp.	Teen Social.	Attitudes Business	Mental Health	Sport Comp.	Teen Social.	Attitudes Business	Mental Health
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Attention	-0.089	0.007	-0.183^{***}	0.029	0.005	0.049	-0.120^{**}	0.003
	[0.055]	[0.052]	[0.065]	[0.058]	[0.048]	[0.048]	[0.060]	[0.045]
Conduct	0.115^{***}	0.080^{*}	-0.007	-0.028	0.138^{***}	0.051	0.033	-0.050
	[0.040]	[0.045]	[0.052]	[0.050]	[0.041]	[0.039]	[0.050]	[0.042]
Emotion	-0.053	-0.093^{**}	0.032	-0.134^{***}	-0.120^{***}	-0.041	-0.001	-0.029
	[0.035]	[0.040]	[0.043]	[0.039]	[0.030]	[0.034]	[0.039]	[0.032]
Peer	-0.205^{***}	-0.134^{***}	0.093^{*}	-0.035	-0.054	-0.127^{***}	0.020	-0.051^{*}
	[0.039]	[0.038]	[0.051]	[0.041]	[0.033]	[0.032]	[0.042]	[0.030]
Cognition	-0.032	-0.028	0.145^{***}	0.034	0.011	-0.110^{***}	-0.095^{*}	0.011
	[0.045]	[0.047]	[0.056]	[0.049]	[0.042]	[0.042]	[0.050]	[0.044]
Family SES	0.072^{*}	-0.084^{**}	0.056	0.013	0.110^{***}	-0.003	-0.057	0.058^{*}
	[0.043]	[0.040]	[0.054]	[0.041]	[0.035]	[0.033]	[0.042]	[0.031]
Backg. controls	X	X	X	X	X	X	X	X
Mean of Dep. Var.	0	0	0	0	0	0	0	0
N individuals	1,095	1,402	775	1,180	1,542	1,975	1,072	1,696

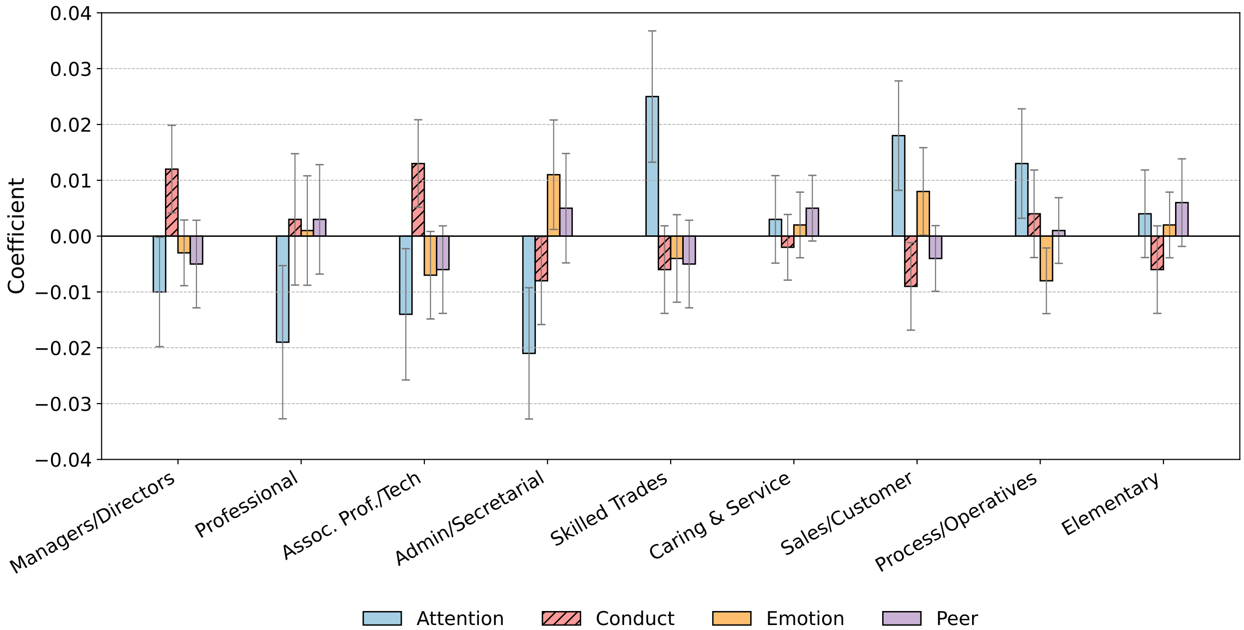
Notes: Data from BCS70. Each column reports measurement-error-corrected estimates from gender-specific regressions of standardised age-16 outcomes (see Table 6) on standardized socio-emotional skills, cognition, and family SES. All specifications include controls for number of siblings, dummies for first child, no dad at birth, teenage mother, year, and region fixed effects. Standard errors in square brackets are estimated from 250 bootstrap replications: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table C.12: Earnings Determinants, Including Career Interests

	[1]	[2]	[3]	[4]	[5]	[6]
Business	0.037***					
	[0.011]					
Practical		0.066***				
		[0.016]				
Living			−0.014			
			[0.010]			
Communication				−0.003		
				[0.012]		
Art					−0.046***	
					[0.011]	
Caring						−0.034***
						[0.012]
Attention	−0.066***	−0.074***	−0.070***	−0.072***	−0.075***	−0.072***
	[0.018]	[0.017]	[0.018]	[0.018]	[0.018]	[0.017]
Conduct	0.051***	0.051***	0.052***	0.053***	0.051***	0.052***
	[0.014]	[0.014]	[0.014]	[0.014]	[0.014]	[0.014]
Emotion	−0.028**	−0.026**	−0.026**	−0.027**	−0.028**	−0.026**
	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]	[0.012]
Peer	−0.005	−0.001	−0.004	−0.003	−0.002	−0.004
	[0.014]	[0.014]	[0.014]	[0.014]	[0.014]	[0.014]
Cognition	0.071***	0.069***	0.074***	0.073***	0.069***	0.069***
	[0.017]	[0.017]	[0.017]	[0.017]	[0.017]	[0.017]
Family SES	0.090***	0.088***	0.089***	0.089***	0.089***	0.088***
	[0.015]	[0.015]	[0.015]	[0.015]	[0.015]	[0.015]
Backg. controls	X	X	X	X	X	X
Mean of Dep. Var.	7.20	7.20	7.20	7.20	7.20	7.20
N individuals	1,847	1,847	1,847	1,847	1,847	1,847
N individual-years	6,729	6,729	6,729	6,729	6,729	6,729

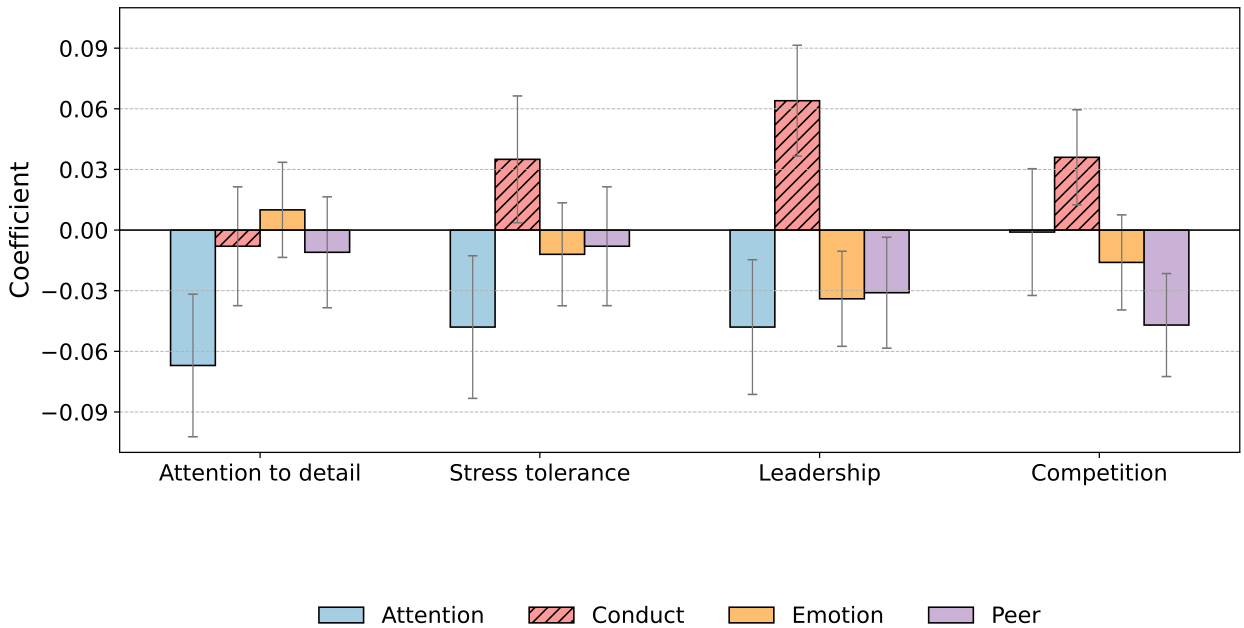
Notes: Data from BCS70. Each column reports measurement-error-corrected estimates from a regression of log real monthly earnings on one of six standardised career interest scores from the JIIG-CAL Questionnaire (1986), together with standardised socio-emotional skills, cognition, and family socio-economic status (SES), constructed using our dedicated measurement system (see equation 1). One career interest variable is introduced at a time. All specifications include controls for gender, gender-by-year, number of siblings, dummies for first child, no father at birth, teenage mother, and year and region fixed effects. Standard errors in square brackets are estimated from 250 bootstrap replications clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Figure C.1: Occupational Sorting by Socio-Emotional Skills



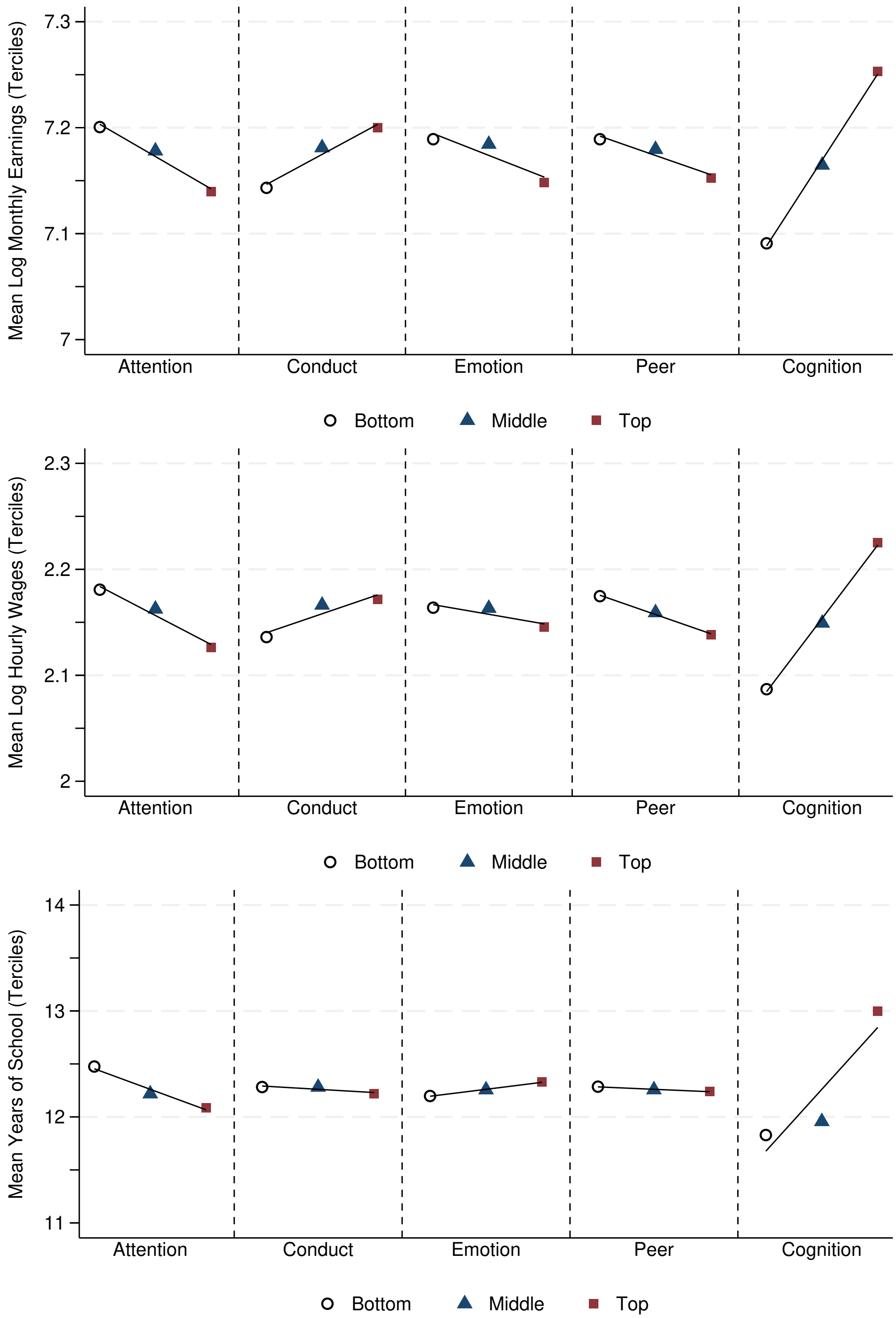
Note: Data from BCS70. The figure shows measurement-error-corrected estimates from a linear probability model (LPM) for 1-digit SOC2010 occupations on socio-emotional skills obtained from our dedicated measurement system (see equation (1)). All specifications control for cognitive ability, socio-economic status, gender, gender-by-year interactions, number of siblings, indicators for first child, no father at birth, teenage mother, year and region fixed effects. Standard errors are estimated from 250 bootstrap replications clustered at the individual level: The figure shows 95% confidence intervals ($p < 0.05$).

Figure C.2: Occupational Requirements (O*NET) by Socio-Emotional Skills



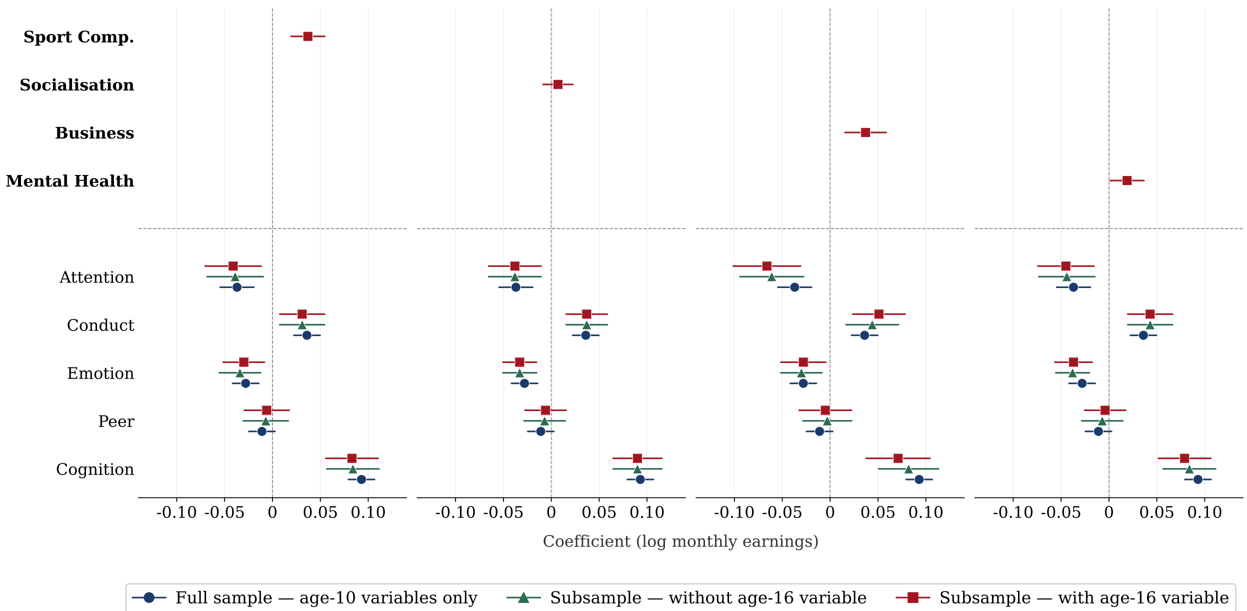
Note: Data from BCS70 and O*NET. The figure shows measurement-error-corrected estimates from a linear probability model (LPM) for standardized O*NET task-importance scales within occupations on socio-emotional skills obtained from our dedicated measurement system (see equation (1)). Occupational codes harmonized to UK SOC 2010 using ONS crosswalks; O*NET data matched using ISCO-08. We use O*NET Likert-type scales for *Work Styles* and *Work Contexts* typically ranging from "Not important at all" to "Extremely important". Work Style: *Attention to detail* indicates whether the job requires being careful about details and thorough in completing tasks; *Stress tolerance* reflects whether the job requires accepting criticism and dealing calmly and effectively with high-stress situations and *Leadership* indicates whether the job requires a willingness to lead, take charge, and provide direction. Work Context: *Competition* measures the extent to which the job requires the worker to compete or be aware of competitive pressures. All specifications control for cognitive ability, socio-economic status, gender, gender-by-year interactions, number of siblings, indicators for first child, no father at birth, teenage mother, as well as year and region fixed effects. Standard errors are estimated from 250 bootstrap replications clustered at the individual level: The figure shows 95% confidence intervals ($p < 0.05$).

Figure C.3: Main Outcomes, by Tercile of Age-10 Skills



Note: Data from BCS70. The figure shows mean log monthly earnings in the top panel, mean log hourly wages in the middle panel and mean years of school in the bottom panel by terciles of socio-emotional skills and cognition. Mean estimates for each socio-emotional skill and cognition are obtained after partialling out: the other socio-emotional skills (cognition), family socio-economic status, gender, gender-by-year, number of siblings, dummies for first child, no dad at birth, teenage mother, year and region fixed effects.

Figure C.4: Earnings Determinants, Including Teen Behaviours



Note: Data from BCS70. The figure plots coefficient estimates and 95% confidence intervals from regressions of log real monthly earnings on standardised socio-emotional skills, cognition, and family socio-economic status (SES). Three sets of estimates are shown: the full sample (navy circle), the age-16 subsample (forest green triangle), and the age-16 subsample augmented with one age-16 variable at a time (red square). All specifications include controls for gender, gender-by-year, number of siblings, dummies for first child, no father at birth, teenage mother, and year and region fixed effects. Standard errors are estimated from 250 bootstrap replications clustered at the individual level: The figure shows 95% confidence intervals ($p < 0.05$).

D Comparison of skills measures with alternative measures of socio-emotional skills and traits

This appendix examines how the labour market effects of our socio-emotional skills measures relate to those for other prominent scales used in the literature.[40] The most widely recognized framework in this area is the Five Factor Model from personality psychology (McCrae & Costa, 1999), commonly known as the 'Big Five'. Alderotti et al. (2023) provide a meta-analysis of labour market returns to skills based on these measures. Further contributions have come from studies using Nordic military assessment data (Edin et al., 2022; Izadi & Tuhkuri, 2024; Jokela et al., 2017).

In terms of how these scales relate to each other, Almlund et al. (2011) and Belfi and Borghans (2025) provide comprehensive discussions of how the Big Five model relates to various other measurement scales. The first purpose of this Appendix is therefore to extend these comparisons and discuss relationships between all of the Big Five, the Nordic military assessments and our factors. To facilitate the discussion, we summarize key relationships between scales in Table D.1, which we refer to throughout. Following this, we relate our work to a parallel literature that examines measures of social engagement more specifically (Deming, 2017; Weinberger, 2014).

D.1 Comparison of age-10 skills measures with Big Five Personality Traits

The Big Five model is a broad personality framework encompassing five traits: openness, conscientiousness, extraversion, agreeableness, and neuroticism (also known as OCEAN traits). Since there is no way to compare these to the factors we derive directly, we rely on existing studies that relate Big Five traits to those from the Strengths and Difficulties Questionnaire (SDQ). The close connection between our measures and the SDQ framework is discussed in Section 5. Henceforth, and for the purposes of this discussion, we consider the SDQ and our factors as interchangeable.

It is important to note that while economists use both Big Five and SDQ-type scales to investigate similar questions, these instruments originate from distinct branches of psychology with different conceptual foundations. The Big Five model originates in personality psychology and adopts a value-neutral approach (i.e. having more extraversion is not considered 'better' in psychological terms than having less). In contrast, the SDQ has its roots in clinical psychology and is designed to capture latent manifestations of psychopathology, where abnormally high scores may indicate potential difficulties requiring intervention.

Nevertheless, several studies have examined the overlap between these two frameworks, following the intuition that problem behaviour in childhood may be an extreme manifestation of

[40] As we emphasize throughout the paper, while the domain of non-cognitive skills has been repeatedly shown to be important in economic outcomes, there is no consensus on the correct representation or even terminology. Almlund et al. (2011) focus on the term 'personality' skills. For convenience, we use the term 'socio-emotional' throughout this appendix.

broader personality traits. For example, the presence of emotional problems might reflect exceptionally high levels of neuroticism. Most studies comparing the two rely on measures collected at the same point in time, often via self-reports from small samples of children or adolescents (Lewis et al., 2014; Muris et al., 2005; Slobodskaya, 2007). A common finding across this literature is that child emotional difficulties explain between a quarter and a third of the variation in personality traits. While there is not a simple one-to-one mapping between SDQ domains and Big Five personality traits, some correlations are notably stronger than others. The first column of Table D.1 summarizes the strongest relationships using evidence from Lewis et al. (2014), where we use our labels for the SDQ domains.

Table D.1: SDQ/Child Behaviours and Other Skill Measures

SDQ/Child Behav. Problems	**Big Five Traits**	**Swedish Military Assessment**	**Finnish Military Assessment**
Attention	↓ Conscientiousness	Negative	↓*Deliberation** ↓ Achievement Striving
Conduct	↓ *Agreeableness** ↑ *Neuroticism** ↑ Extraversion	Positive (through ↓ *Agreeableness** ↑ Extraversion)	↑ Sociability ↓ Dutifulness
Emotional	↑ *Neuroticism** (strong) ↓ Extraversion	Negative	↓ Self-confidence ↓ *Deliberation**
Peer	↓ Extraversion ↑ *Neuroticism** ↓ *Agreeableness**	Negative	↓ Sociability ↓ Leadership Motivation

Key: ↑ positive association; ↓ negative association; (strong) particularly strong relationship. Italics and asterisk indicate an inverse association with earnings (e.g., the effect of emotional problems on earnings through higher neuroticism is negative; but the effect of conduct problems through lower agreeableness is positive).
Notes and Sources: Big Five associations refer to partial correlations based on Lewis et al. (2014); Swedish Military Assessment relationships are based on Edin et al. (2022); Finnish Military Assessment relationships are based on Jokela et al. (2017). See text for details.

We attempt to relate our findings on the SDQ-type measures quantitatively to those from the Big Five using the following exercise. We predict the labour market returns to SDQ behaviours by combining: i) their relationship with the Big Five traits and ii) the established returns to the Big Five traits. The idea is that an alignment of these predicted returns with our main findings shows a consistency of our findings with the literature on returns to the Big Five.

With this in mind, the LM column of Table D.2 shows returns to Big Five traits from Alderotti et al. (2023)'s meta-analysis of approximately 90 studies, which provides the most recent evidence on returns across multiple populations. All traits have statistically significant relationships with earnings, with the strongest relationship being for agreeableness (where higher agreeableness leads to lower earnings) and the weakest for openness.

Table D.2: Implied Returns to SDQ Behaviours Through Big-Five Traits

Big-Five traits	**LM Returns (Alderotti et al., Tbl 3)**	**Correlations (Lewis et al., Tbl 2)**			
		Attention	Conduct	Emotion	Peer
Neuroticism	-0.032 [0.0054]	0.25	0.29	0.51	0.20
Extraversion	0.021 [0.0043]	0.01	0.04	-0.26	-0.34
Openness	0.018 [0.0073]	0.05	-0.01	0.06	0.01
Agreeableness	-0.035 [0.0052]	-0.13	-0.26	0.06	-0.13
Conscientiousness	0.025 [0.0043]	-0.33	-0.23	-0.06	-0.09
Implied returns:		-0.006	0.001	-0.018	-0.011

Sources: Alderotti et al. (2023), Lewis et al. (2014).
Notes: Standard errors on labour-market returns from Alderotti et al. (2023) are shown in square brackets below the point estimates. Implied returns are computed using the coefficients from regressions of Big-Five traits on SDQ skills as described in the text. SDQ labels correspond to our convention; Lewis et al. denote them as "Hyperactivity" (Attention), "Conduct", "Anxiety" (Emotion), and "Peer".

The remaining columns of Table D.2 show correlations between Big Five traits and SDQ-measured skills from Lewis et al. (2014), who collect self-reports from 16-year-olds. We use this covariance matrix to calculate regression coefficients for each Big Five trait on the four SDQ measures.[41] Specifically, for each Big Five trait i, we compute:

$$\boldsymbol{\beta}_i = \mathbf{R}_{SDQ}^{-1}\mathbf{r}_i$$

where $\mathbf{R}_{SDQ}$ is the 4×4 covariance matrix among the SDQ/child behaviour measures (attention, conduct, emotional, and peer problems), shown in Table B.10, $\mathbf{r}_i$ is the 4×1 vector of covariances between each SDQ measure and Big Five trait i, shown in the right-hand columns of Table D.2, and $\boldsymbol{\beta}_i$ is the resulting 4×1 vector of regression coefficients. These regression coefficients can be interpreted as the partial effect of changes in SDQ behaviours on later personality trait i. We do not show them here, but in terms of directions of effects, they closely match those in $\mathbf{r}_i$.

The predicted labour market return to SDQ measure j through associations with Big Five traits is then calculated as:

[41] Data from Lewis et al.'s Table 2. As all variables are standardized we use the terms correlation and covariance interchangeably.

$$\pi^{predicted}_{SDQ,j} = \sum_{i=1}^{5} \beta_{j,i} \cdot \pi_{Big5,i} \tag{6}$$

where $\beta_{j,i}$ is the regression coefficient of Big Five trait i on SDQ measure j (controlling for other SDQ measures), and $\pi_{Big5,i}$ is the labour market return to Big Five trait i, shown on the left hand side of Table D.2.

The bottom row of Table D.2 shows these predicted returns. The strongest effects emerge for the internalising behaviours: a standard deviation increase in emotional problems is here predicted to reduce earnings by 0.018 log points, driven by strong positive relationships with neuroticism and negative relationships with extraversion, while peer problems show a slightly weaker predicted association at -0.011. Although we do not formally calculate standard errors, we note that the correlation coefficients for the skills measures are highly precise (Lewis et al. report that correlations of 0.06 are significant at the 5% level). Using the standard errors on the returns to Big Five traits as a guide, we expect that the implied returns on emotional problems are strongly significant and the returns on peer problems are borderline significant at the 5% level.

Conduct problems are predicted to have essentially zero returns (0.001). This occurs because offsetting relationships cancel each other out: while conduct problems predict higher neuroticism and lower conscientiousness (both implying negative returns), they also predict higher extraversion and lower agreeableness, patterns associated with positive earnings in the labour market.

Our takeaway from this exercise is that the extensive research on Big Five traits is consistent with conduct problems having non-negative labour market returns. Our exercise can also be interpreted as a mediation analysis, where we calculate the returns to childhood SDQ behaviours that operate through their influence on adult personality traits. In this light, our findings extend beyond this mediation by documenting significantly *positive* returns to conduct behaviours. Our analysis therefore adds to what would have been predicted from examining the Big Five literature alone, and shows that the returns to conduct behaviours are not fully explained through their effect on later Big Five traits.

D.2 Comparison of age-10 skills measures with Nordic armed forces measures

We now discuss the relationship of our skills measures with those taken from Swedish and Finnish military assessments, taken at age 18 for those entering the armed services. These assessments provide valuable information on skill measures that can be linked to administrative data on labour market outcomes. It is important to note, however, that they are representative of males only.

D.2.1 Swedish Military Assessments

Swedish studies (e.g. Edin et al., 2022; Lindqvist & Vestman, 2011) derive a measure of broad non-cognitive skills based on semi-structured interviews conducted by certified psychologists. Interviews cover five key areas: school experience, work experience, leisure activities (with particular focus on leadership roles), home environment/upbringing, and emotional stability. The psychologists rate conscripts on a standardized scale using these interviews as well as reviewing various other information sources including cognitive test results, school grades, physical tests, and questionnaires about family, friends, and hobbies. These assessments capture traits such as social maturity, psychological energy, intensity, and emotional stability (Lindqvist & Vestman, 2011).

In the absence of any literature documenting how this composite measure of non-cognitive skills relates to SDQ-type behaviours, we assess these relationships by using associations with the Big Five. Edin et al. (2022) investigate how the non-cognitive skill measure relates to the Big Five by correlating their skill measures with job requirements given by O*NET job descriptors. The partial correlations are presented in Table D.3, reproduced from Edin et al. (2022) Table A.1. The strongest associations found are for individuals with high non-cognitive skill (controlling for cognitive skills) to sort into occupations requiring greater extraversion and lower agreeableness. Weaker correlations are found for conscientiousness (positive) and for neuroticism (negative).

We can then use this information to assess how the Swedish measured non-cognitive skill relates to our SDQ-type skills. We adapt (6) to generate the following coefficients:

$$\phi_{SDQ,j}^{predicted} = \sum_{i=1}^{5} \beta_{j,i} \cdot \phi_{Big5,i}$$

where $\phi_{Big5,i}$ is the partial association of Big Five trait i with the Swedish 'non-cognitive' measure, shown on the left hand side of Table D.3, and $\beta_{j,i}$ is the partial effect of SDQ behaviour j on Big Five trait i, computed above. Then $\phi_{SDQ,j}$ captures the resulting partial association of SDQ behaviour j on the non-cognitive measure. Similarly to before we can interpret the predicted effects as the effects of child behaviours on the non-cognitive measure that is mediated through Big Five traits.

The results are shown on the right-hand side of Table D.3. They show negative predicted effects of all the child behavioural problems except for conduct. This positive result for conduct is driven by a positive relationship with extraversion, which is positively associated with non-cognitive skill, and the negative relationship with agreeableness which in turn is negatively associated with non-cognitive skill. In this case, given the precision of the underlying estimates presented on the left-hand side, we conjecture that the right-hand side coefficients are significant. Again the key associations are summarized in the middle column of Table D.1.

Table D.3: SDQ Behaviours and Military-Assessed Non-Cognitive Skills

Big-Five traits	Rel. to Non-Cog Skill (Edin et al., Tbl A.1)	Child Behaviours	Implied Effect on Non-Cog Skill
Neuroticism	-0.0990 [0.0036]	Attention	-0.021
Extraversion	0.1863 [0.0032]	Conduct	0.045
Openness	-0.0434 [0.0033]	Emotion	-0.073
Agreeableness	-0.1695 [0.0043]	Peer	-0.062
Conscientiousness	0.0989 [0.0042]		

Sources: Edin et al. (2022) and Lewis et al. (2014), based on analysis described in the text.
Notes: Standard errors on Big-Five coefficients from Edin et al. (2022) shown in square brackets. Implied effects are computed by combining (i) the partial associations of Big Five traits with the Swedish non-cognitive measure (left side of table) and (ii) regressions of Big Five traits on SDQ behaviours, as described above.

D.2.2 Finnish Military Assessments

Studies using Finnish military data (Jokela et al., 2017; Izadi & Tuhkuri, 2024) use a personality test that captures eight distinct traits: sociability, activity-energy, masculinity, dutifulness, deliberation, achievement striving, leadership motivation, and self-confidence. These are derived from 218 yes/no individual statements, with each trait formed by summing scores for 18-33 items. In a regression controlling for all traits as well as cognitive skills, Jokela et al show that all traits are positively associated with earnings except for deliberation and masculinity.

Similarly to before, in the absence of direct evidence relating these measures to SDQ-type behaviours, we rely on quantifying the effects of SDQ behaviours that are mediated through Big Five traits. On this front, we make use of results from an auxiliary survey that Jokela et al. administer, which elicits both the Big Five traits and the Finnish military traits (through a shortened questionnaire). We then compute a quantitative measure of the relationship between the SDQ behaviours and the Finnish military assessments by combining the partial correlations Jokela et al. obtain alongside the partial associations of the SDQ behaviours with Big Five traits computed above. Specifically, we now adapt (6) as follows:

$$\psi_{SDQ,j}^{k,predicted} = \sum_{i=1}^{5} \beta_{j,i} \cdot \psi_{Big5,i}^{k}$$

where, on the right hand side, $\psi_{Big5,i}^{k}$ is the partial association of Big Five trait i with Finnish measure k (e.g. sociability), shown in Jokela et al.'s Table S.15 and not reproduced here, and $\beta_{j,i}$ is defined as before. $\psi^{k}_{SDQ,j}$ then captures the implied partial association of SDQ behaviour j on Finnish measure k.

The results for $\psi^{k,predicted}_{SDQ,j}$ are shown in Table D.4. Conduct problems most strongly affect sociability (positively) and dutifulness (negatively), while emotional problems reduce self-confidence, and also heighten deliberation. As mentioned previously, deliberation has a negative labour market return (Jokela et al.'s Table S.2).

Table D.4: Effect of SDQ Behaviours on Finnish-Assessed Skills

	Leadership Motivation	**Activity-Energy**	**Achievement Striving**	**Self-Confidence**	***Deliberation (-ve return)***	**Sociability**	**Dutifulness**	***Masculinity (-ve return)***
Attention	−0.01	−0.07	−0.14	−0.04	−0.18	0.07	−0.09	−0.04
Conduct	0.05	0.01	−0.03	−0.05	−0.04	0.10	−0.11	0.00
Emotion	−0.12	−0.11	0.08	−0.24	0.19	−0.16	0.00	−0.05
Peer	−0.21	−0.16	−0.08	−0.14	0.13	−0.25	0.00	−0.01

Sources: Based on Jokela et al. (2017) Table S.15 and Lewis et al. (2014) Table 2, as described in the text.
Notes: Table shows estimated effects of a one standard deviation increase in each SDQ behaviour problem on personality traits measured in Finnish armed-forces assessments. Deliberation and Masculinity are highlighted as traits with negative labour market returns (see Jokela et al. (2017) Table S.2).

We can go one step further and, similarly to before, we can impute labour market returns by weighting with the returns on the Finnish measures. Although not shown here, the implied return for conduct is essentially zero. Returns for attention and emotion problems are mildly negative, while the strongest predicted negative return is for peer problems. This is largely generated by the strong negative effects for peer problems on Leadership Motivation and Sociability. Similarly to above, the key associations are summarized in the final column of Table D.1.

D.3 Comparison of age-16 measures with social skills measures in the literature

Our measures of sport competition and teen socialization at age 16 discussed in Section 6.4 and presented in Table 6 overlap with social skills measures constructed by Deming (2017) and Weinberger (2014). This section examines similarities and differences, and compares our findings with theirs.

D.3.1 Comparison of measures

Our 'teen socialization' measure broadly captures the extent of interacting with friends outside school, including activities such as spending time at friends' homes, having friends visit one's home, going out with friends, and socializing in various contexts during school term. Our 'sports competition' measure captures the number of sports in which teens represented their school or club in the previous year.

From the NLSY79, Deming constructs a composite measure of social skills using self-reported sociability in 1981, retrospective self-reported sociability at age 6, the number of clubs in which the respondent participated in high school, and participation in high school sports. These were elicited when respondents were aged 16-24. When performing a comparison across cohorts, Deming uses only the sociability measures (dropping participation in clubs and sports) and, from the NLSY97, two items from the Big Five inventory capturing extraversion.

Weinberger uses participation in extracurricular activities including: individual and team sports, clubs (academic and other), performing arts, student publications, and leadership roles in various clubs and societies. Her measures particularly emphasize leadership positions within these activities.

Broadly speaking, our measure of teen socialization overlaps with Deming's measure of sociability, albeit with a slight difference in timing. Our measure of sports competition overlaps with Deming's measure of participation, although ours is more focused on competitive aspects. The sports competition measure also overlaps with Weinberger's construct, though hers focuses more on leadership roles and explicitly outward-facing activities (like performing arts).

An important limitation of participation measures is the availability of the activity itself: participation in clubs presumably partly reflects accessibility and local culture, which may capture family, schooling, and local area privileges. This may be slightly captured in our sports competition measure, but is arguably more reflected in Deming's broader measure and even more in Weinberger's measures. This aspect could confound the relationship between these measures and later outcomes. In contrast, as discussed in Section 6.3, in terms of our age-10 skills we show that when we control for family or other environmental factors shared by siblings, a large part of the estimated returns to these skills remains.

D.3.2 Comparison of findings

Returns: To compare our findings with these earlier papers, we examine the relationship between our age-16 measures and earnings, with and without controls for our age-10 skills. Table D.5 presents these results.

These results show that the returns to sport competition are much stronger than for socialization, which by itself has an insignificant relationship with earnings. As referenced in Section 6.4, part of this effect of socialization operates through reduced education, and although not shown here, when education is included in controls, the direct effect of socialization on earnings is significantly positive.

More importantly for the discussion, these results show that controlling for age-10 skills has a small impact on the estimated effects of teen socialization and sports competition, suggesting that these age-16 measures capture partly independent aspects of human capital. Our two age-16 mea-

Table D.5: Earnings Effects of Age-16 Behaviours

	Earnings			
	[1]	[2]	[3]	[4]
Sport Competition	0.044***	0.037***		
	[0.009]	[0.009]		
Socialization			0.012	0.007
			[0.008]	[0.008]
Attention		−0.041***		−0.038***
		[0.015]		[0.014]
Conduct		0.031**		0.037***
		[0.012]		[0.011]
Emotion		−0.030***		−0.033***
		[0.011]		[0.009]
Peer		−0.006		−0.006
		[0.012]		[0.011]
Cognition	0.112***	0.083***	0.118***	0.090***
	[0.011]	[0.014]	[0.010]	[0.013]
Family SES	0.087***	0.089***	0.082***	0.083***
	[0.011]	[0.012]	[0.010]	[0.010]
Backg. controls	X	X	X	X
Mean of Dep. Var.	7.18	7.18	7.18	7.18
N individuals	2,637	2,637	3,377	3,377
N individual-years	9,621	9,621	12,232	12,232

Notes: Data from BCS70. Each column reports measurement-error-corrected estimates from a regression of log monthly earnings on standardised age-16 measures (sport competition in columns [1]–[2] and socialization in columns [3]–[4]), together with age-10 socio-emotional skills, cognition, and family socio-economic status (SES), constructed using our dedicated measurement system (see equation 1). All specifications include controls for gender, gender-by-year, number of siblings, dummies for first child, no father at birth, teenage mother, and year and region fixed effects. Standard errors in square brackets are estimated from 250 bootstrap replications clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

sures map onto the distinction Deming (2017) draws between sociability and broader participation: teen socialization aligns with his sociability-only composite, while our sports competition measure aligns with his participation-inclusive composite, which incorporates involvement in clubs and sports. Consistent with this mapping, we find a substantially stronger association between earnings and sports competition (0.037) than between earnings and teen socialization (0.007), echoing the pattern in Deming's own estimates: his sociability-only coefficient on log hourly wages is 0.020 (Table IV, Column 4), compared with 0.043 for his participation-inclusive composite (Table I, Column 4). We compare to Column 4 in each of Deming's tables as this excludes years of education, matching our baseline specification.

Complementarity: One of the key takeaways from Deming's and Weinberger's results is complementarity between cognitive and social skills. As referenced in Section 6.2, and shown in Table C.8, we find little evidence of complementarity between cognition and our age-10 socio-emotional skills measures.[42] We offer an explanation for this discrepancy. When using only sociability measures (without participation components) in cross-cohort comparisons, Deming finds no significant interaction between cognitive and social skills (his Table IV). The interactions seem to matter only when incorporating participatory activities in the measures of social skills (his Table I). Similarly, Weinberger focuses exclusively on measures of participation and leadership and finds strong evidence of complementarity between these and cognitive skills. This suggests that the complementarity identified in the these earlier papers may be particularly associated with the leadership and participation aspects of social skills rather than with the social competence and regulation measures we examine here.

Discussion: Overall, our takeaways from this analysis are as follows: First, the persistence of the effects of teen socialization and sports competition even after controlling for our four socio-emotional factors suggests these activities capture dimensions of human capital development that only partially overlap with our measures of childhood skills. This finding helps contextualize the earlier social skills literature, particularly the two prominent papers cited here. Second, we find no strong evidence of complementarity between cognitive and social skills here, indicating that the complementarity found particularly by Weinberger may be more a feature of social skills as measured by participatory and leadership activities. Third, as discussed in Section 6.3, we find that even when controlling for shared parental and environmental factors, a large share of our estimated returns to age-10 socio-emotional skills remains. This may be less the case for age-16 measures of social skills, which may be confounded by family, school, or local area effects. Unfortunately the lack of information on school characteristics at age 16 prohibits testing this hypothesis cleanly.

In summary, the absence of complementarity between cognitive skills and our relevant socio-emotional measures, in contrast to findings by Deming (2017) and Weinberger (2014), possibly reflects basic differences in what these measures capture. Our teacher-assessed behaviours at age 10 represent core socio-emotional skills, while the participation and leadership measures in the literature capture more complex social experiences that may particularly benefit high-cognitive individuals. This distinction suggests that fundamental behavioural skills and participatory social activities represent partly independent aspects of human capital development.

[42] We find a significant interaction between cognition and attention, which is consistent with attention operating in part through the same schooling channel as cognition itself, and distinct from the social-skills complementarity at issues in Deming and Weinberger.

E Further Details on Use of NLSY Data

Data

The NLSY79 Child and Young Adult Survey (NLSY79 CYA) tracks the children of women in the original National Longitudinal Survey of Youth 1979 cohort, surveyed biennially since 1986. Our analysis includes 6,028 children born between 1971 and 1994 with at least one labour market observation between 2000 and 2020, the full available window, and complete data on childhood cognitive and socio-emotional measures. Because the NLSY79 CYA follows *all* children of original NLSY79 women, the data allows for a sibling fixed-effect design.

We construct real monthly earnings by using information on annual earned income (e.g., wages, salary) deflated by the Consumer Price Index (CPI, base year 2015). We winsorize earnings at the 3rd and 97th percentiles within age groups. Years of schooling are estimated using information on the highest grade of formal education and the highest academic degree reported before age 30.[43]

Socio-economic and demographic characteristics, including sex, birth date, ethnicity, maternal education (grade level and degree), and birth order, were obtained from cross-wave variables in the mother's 1979 screener. Whether the NLSY79 mother gave birth as a teenager is identified using the birth date of her first child. Real family income (a measure of background SES) is derived from total net family income reported by the parents in the main NLSY79 surveys before respondents reach age 30, deflated using the CPI (base years 1982–84 = 100).[44] Biological sibling pairs are identified using kinship links from the NLSY79 and NLSY79 CYA datasets (Rodgers et al., 2016).[45] Table E.1 compares key statistics between NLSY79 CYA and BCS70, and conducts tests of differences across samples.

We measure socio-emotional skills using the Behavior Problems Index (BPI) which is used to assess the range and type of childhood problems of behaviour for children over the age of four. The BPI is constructed from a set of 32 questions answered by parents and teachers about child behaviour, with responses given as 'often', 'sometimes', and 'never'.[46] Rather than present the full list of items we refer the reader to other studies using these data such as Reyes (2015).

An exploratory factor analysis suggests that these items load onto two broad dimensions: externalising and internalising behaviours. To construct a four-factor skill representation in the NLSY79 CYA comparable to that used in our main analysis, we follow Lise and Postel-Vinay (2020), who apply principal component analysis (PCA) with exclusion restrictions to construct skill measures.

[43] Missing values are filled using the first available report from subsequent survey waves.

[44] Approximately 85% of respondents were age 29 or 30 at the time of reporting. We exclude individuals whose parents reported income only before age 21 or after age 30 (about 0.8% of the sample).

[45] The data are publicly available at doi:10.1007/s10519-016-9785-3.

[46] We use the BPI collected around age 10 to facilitate comparability with the BCS70 socio-emotional constructs. For approximately 85% of the sample, measures are observed between ages 9 and 11. For the remaining observations, we rely on measures collected at slightly younger ages (around 13%) or during adolescence (around 2%).

Table E.1: NLSY79 CYA vs BCS70 Summary Statistics: Relevant Characteristics

	BCS70		NLSY79 CYA		Test	p-val
	Obs	**Mean**	**Obs**	**Mean**		
	[1]	[2]	[3]	[4]	Stat. (diff. / χ^2)	[5]
					Mean diff.	
Year	23,451	2005.7	30,583	2010.1	-4.43	0.00
		[6.52]		[5.71]		
Age	23,451	35.7	30,583	25.9	9.81	0.00
		[6.52]		[5.97]		
Years of Schooling	6,952	12.26	6,028	13.12	-0.86	0.00
		[2.31]		[2.08]		
Parent's highest qualification	6,952	2.05	6,028	2.86	-0.81	0.00
		[1.27]		[1.32]		
Number of siblings	6,952	1.58	6,028	2.93	-1.35	0.00
		[1.04]		[1.81]		
Binary variables					**Pearson χ^2**	
Female	6,952	0.51	6,028	0.50	4.15	0.042
		[0.50]		[0.50]		
First child	6,952	0.42	6,028	0.46	23.81	0.00
		[0.49]		[0.50]		
Teenage mother	6,952	0.08	6,028	0.18	251.82	0.00
		[0.28]		[0.38]		

Data from BCS70 and NLSY79 Child and Young Adult cohort. This table compares sample characteristics between the BCS70 and the NLSY79 CYA. Parental education is mapped to the BCS70 categorization using the mother's years of schooling from the NLSY79 CYA dataset. The mapping is as follows: 0–11 years = no qualifications; 12 years = O level; 13 years = A level; 14–15 years = SRN/Certificate in Education; 16+ years = degree level. Means and standard deviations (in brackets) are reported for continuous variables. For binary variables, means correspond to proportions, and standard deviations are also shown in brackets. Column [5] reports p-values for t-tests (continuous variables) and Pearson's χ^2 tests (binary variables) comparing the two samples.

Let $\mathbf{M}_{N\times B}$ denote the matrix of N individuals and B behavioural measures. PCA decomposes $\mathbf{M}_{N\times B}$ as $\mathbf{M} = \mathbf{FL}$, where $\mathbf{L}_{B\times B}$ is the loadings matrix whose rows are the principal eigenvectors of $\mathbf{M}^\top\mathbf{M}$ and $\mathbf{F}_{N\times B}$ is the corresponding matrix of component scores. Retaining only the first four components yields the approximation $\mathbf{M} \approx \mathbf{F}_{N\times 4}\mathbf{L}_{4\times B} + \mathbf{U}$. For any invertible matrix $\mathbf{T}_{4\times 4}$, we can write $\mathbf{M} = (\mathbf{F}_{N\times 4}\mathbf{T}_{4\times 4})(\mathbf{T}^{-1}_{4\times 4}\mathbf{L}_{4\times B}) + \mathbf{U}$, which defines an equivalent decomposition into linearly transformed factors $\mathbf{F}_{N\times 4}\mathbf{T}_{4\times 4}$ with corresponding loadings $\mathbf{T}^{-1}_{4\times 4}\mathbf{L}_{4\times B}$.

We set $\mathbf{T} = \mathbf{L}_{4\times 4}$, the submatrix formed by the first four columns of $\mathbf{L}_{4\times B}$ to satisfy our exclusion restrictions on four anchor items, one per factor, requiring each to load only on its intended factor: *attention* (Has difficulty concentrating/paying attention), *conduct* (Argues too much), *emotion* (Worries too much), and *peer* problems (Has trouble getting along with other children). This pins down the rotation **T** using a single item per factor, while allowing the remaining 28 BPI items to load freely across all four components. These anchors are conceptually similar to the items with the highest loadings in the corresponding dimensions of the BCS70 measurement system. Table E.2 reports the transformed component loadings (transposed for ease of presentation).[47]

Table E.2: Transformed Component Loadings from PCA of 32 BPI Items (NLSY79 CYA)

Variable	**Attention**	**Conduct**	**Emotion**	**Peer**
NLSY79 CYA Domains	**Hyperactive**	**Headstrong**	**Anxious/Depressed & Dependent**	**Peer Problems & Antisocial**
Has difficulty concentrating/paying attention	**1.00**	0.00	0.00	0.00
Argues too much	0.00	**1.00**	0.00	0.00
Worries too much	0.00	0.00	**1.00**	0.00
Has trouble getting along with other children	0.00	0.00	0.00	**1.00**
Is restless, overly active, cannot sit still	**0.94**	0.23	-0.11	-0.10
Is easily confused, seems in a fog	**0.84**	-0.16	0.34	0.04
Clings to adults	**0.62**	0.04	0.58	-0.40
Is too dependent on others	**0.59**	-0.02	0.62	-0.16
Is impulsive or acts without thinking	**0.54**	0.29	-0.06	0.26
Is stubborn, sullen, or irritable	0.07	**0.68**	0.07	0.27
Has strong temper and loses it easily	-0.01	**0.65**	0.01	0.40
Has sudden changes in mood or feeling	-0.06	**0.60**	0.36	0.18
Is disobedient at home	0.05	**0.59**	-0.22	0.48
Demands a lot of attention	0.51	**0.55**	0.38	-0.38
Is too fearful or anxious	0.34	0.17	**0.76**	-0.15
Is unhappy, sad, or depressed	-0.17	0.03	**0.70**	0.62
Feels worthless or inferior	-0.19	0.03	**0.70**	0.61
Cries too much	0.06	0.38	**0.57**	-0.04
Feels/complains no one loves him/her	-0.34	0.48	**0.50**	0.40
Has trouble getting mind off certain thoughts	0.48	-0.06	**0.50**	0.15
Is rather high strung, tense, and nervous	0.32	0.38	**0.48**	-0.02
Has trouble getting along with teachers	0.38	-0.30	-0.27	**1.00**
Is not liked by other children	0.05	-0.27	0.23	**0.93**
Bullies or is cruel/mean to others	-0.08	0.24	-0.16	**0.93**
Is disobedient at school	0.50	-0.27	-0.47	**0.92**
Hangs around with kids who get into trouble	0.36	-0.44	-0.10	**0.92**
Feels others are out to get him/her	0.00	-0.19	0.49	**0.82**
Is withdrawn, does not get involved with others	0.00	-0.49	0.67	**0.80**
Breaks things deliberately	0.20	0.08	0.02	**0.73**
Does not seem to feel sorry after misbehaving	0.14	0.06	-0.10	**0.65**
Is secretive, keeps things to self	0.21	-0.38	0.44	**0.57**
Cheats or tells lies	0.40	0.20	-0.24	**0.50**

Notes: Data from the NLSY79 Child and Young Adult cohort. The table presents transformed loadings from a principal component analysis (PCA) of 32 items from the Behaviour Problems Index (BPI). The raw eigenvector matrix has been rescaled to impose exclusion restrictions such that the first four items: *has difficulty concentrating/paying attention*, *argues too much*, *worries too much*, and *has trouble getting along with other children*, have unit loadings on their respective components. These items serve as anchors for the first four components (explaining 57% of the total variance), representing attention, conduct, emotional, and peer problems respectively.

[47] BPI items are recorded as ordered categorical responses so we estimate the PCA using a polychoric correlation matrix.

The NLSY79 CYA Survey includes subscales of the BPI covering Hyperactivity, Headstrong behaviour, Antisocial behaviour, Anxious/Depressed, Dependent, and Peer Problems. The first and second principal components from the PCA, labelled *attention* and *conduct*, load most strongly on the Hyperactivity and Headstrong subscales, respectively. The third component reflects a combination of Anxious/Depressed and Dependent subscales, which we interpret as an *emotion* construct. The fourth component aligns with Peer Problems in both studies, although in the NLSY79 CYA it also captures elements of antisocial behaviour.

Cognitive skills are derived using a PCA of Peabody Individual Achievement Test (PIAT) scores collected during childhood. We extract the first component from tests assessing mathematical reasoning and problem-solving, reading recognition and comprehension, and pictured vocabulary and verbal ability.

Table E.3 reports the correlations between the resulting socio-emotional and cognitive constructs as well as other variables of interest, and compares them with those obtained in the BCS70.[48]

Table E.3: Correlation of Key Variables: NLSY79 CYA vs BCS70

Dataset	Measure	Atte	Cond	Emot	Peer	Cogn	Fam.SES	Yrs.Sch
NLSY79 CYA	Attention	1.00						
	Conduct	0.55	1.00					
	Emotion	0.31	0.32	1.00				
	Peer	0.52	0.68	0.30	1.00			
	Cognition	-0.30	-0.10	-0.07	-0.18	1.00		
	Family SES	-0.20	-0.09	-0.07	-0.15	0.34	1.00	
	Years School	-0.27	-0.17	-0.04	-0.23	0.43	0.32	1.00
BCS70	Attention	1.00						
	Conduct	0.59	1.00					
	Emotion	0.37	0.27	1.00				
	Peer	0.46	0.37	0.44	1.00			
	Cognition	-0.50	-0.21	-0.23	-0.24	1.00		
	Family SES	-0.19	-0.11	-0.09	-0.10	0.39	1.00	
	Years School	-0.31	-0.17	-0.12	-0.14	0.46	0.38	1.00

Data from BCS70 and NLSY79 CYA cohort. This table reports pairwise correlations among standardized scores representing socio-emotional problems, cognitive ability, family socioeconomic status (SES), and years of schooling in the NLSY79 CYA and the BCS70 cohort. In the NLSY79 CYA, behavioral and cognitive scores are derived from principal component analysis (see Table E.2 for details). Family SES is based on standardized net family income reported in the main NLSY79 surveys before the respondent reaches age 30. Years of schooling reflects the highest grade completed or academic degree attained before age 30.

[48] BPI items are originally recorded such that lower values indicate more problematic behaviour, so we reverse the final scores such that higher values reflect higher behavioural difficulties, consistent with our BCS70 socio-emotional constructs.

F Further Details on Quantification of Intergenerational Transmission of Economic Status

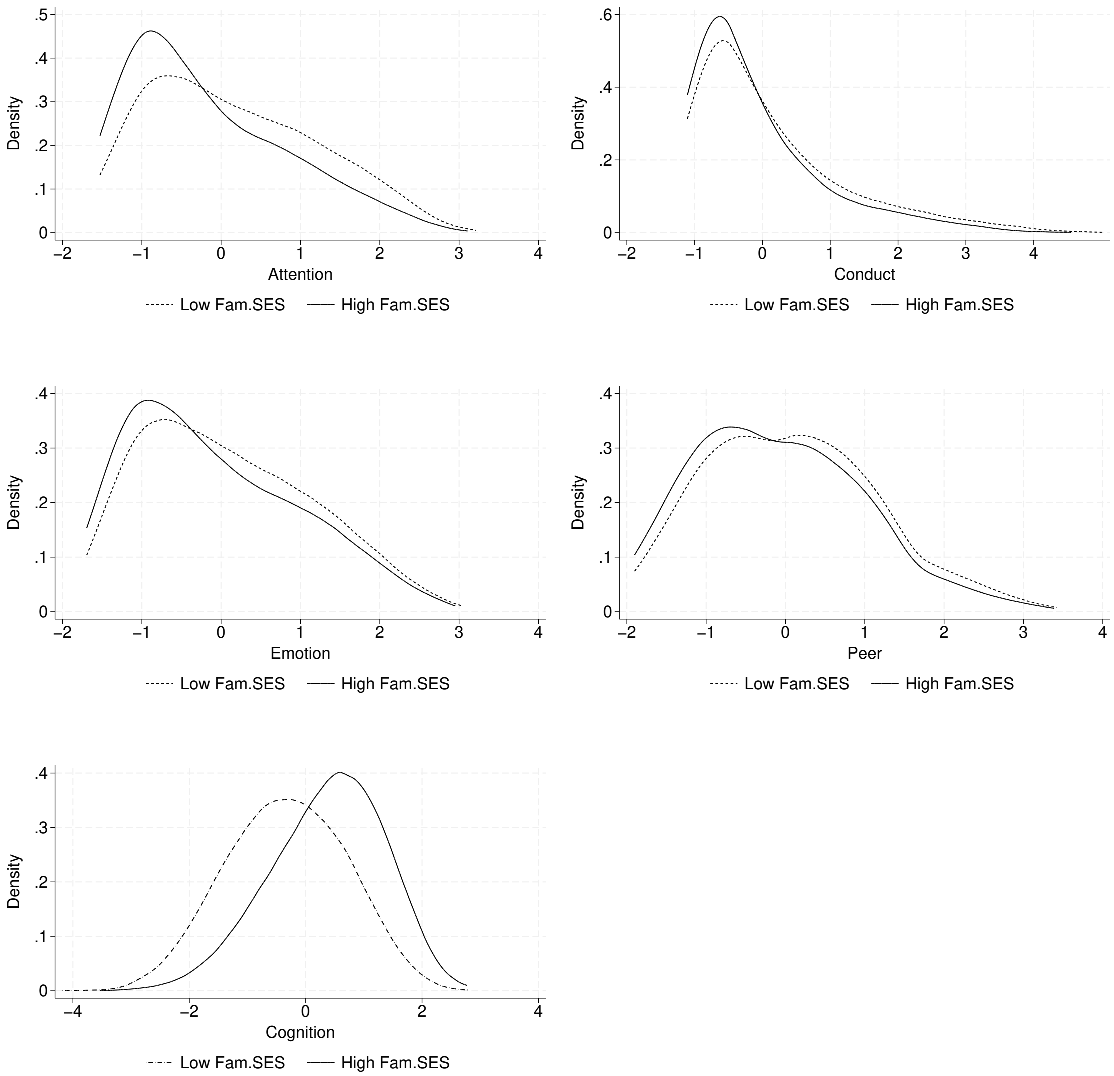
F.1 Results Discussed in Section 7

Table F.1: Differences in Endowments by Family SES

	Group Means		FSD Test	
			Stat	p-val
Panel A: Top vs. Bottom Half Family SES	*Bottom half*	*Top half*		
Attention	0.160	-0.177	0.000	1.00
	[1.050]	[0.991]		
Conduct	0.096	-0.107	0.000	1.00
	[1.094]	[0.948]		
Emotion	0.081	-0.090	0.000	1.00
	[1.056]	[1.062]		
Peer	0.080	-0.089	0.027	0.99
	[1.046]	[1.032]		
Cognition	-0.335	0.372	0.011	1.00
	[1.012]	[0.958]		
N individuals	3,658	3,294		
Panel B: College Degree vs. No Degree	*None*	*Degree*		
Attention	0.063	-0.360	0.014	0.99
	[1.037]	[0.951]		
Conduct	0.031	-0.180	0.000	1.00
	[1.046]	[0.927]		
Emotion	0.030	-0.174	0.000	1.00
	[1.057]	[1.080]		
Peer	0.025	-0.143	0.091	0.96
	[1.043]	[1.032]		
Cognition	-0.131	0.751	0.025	0.99
	[1.019]	[0.884]		
N individuals	5,922	1,030		

Notes: Data from BCS70. The first two columns report group means and standard deviations by two splits of family socio-economic status. Columns three and four show results of the first-order stochastic dominance test proposed by Barrett and Donald (2003). The null hypothesis is H_0: high-SES (College Degree) stochastically dominates low-SES (No Degree), with socio-emotional skills reverse coded. The test statistic in column three is $\sqrt{\frac{NM}{N+M}}\,\sup_x\left(\hat{F}_H(x) - \hat{F}_L(x)\right)$, where $\hat{F}_i(x)$ is the empirical CDF for group i, N and M are the sample sizes, and x denotes the level of skill across the distribution. Socio-emotional constructs are re-oriented so that a higher values indicate fewer problems/higher skills. The distribution of the test statistic and associated p-values are obtained from 1000 bootstrap replications.

Figure F.1: Differences in Distributions of Skills by Family SES



Notes: Data from BCS70. The figure shows kernel density distributions of socio-emotional and cognitive skills, comparing children from low versus high family socio-economic status, defined as the bottom and top halves of the family SES distribution score.

Table F.2: Determinants of Schooling & Earnings by Family SES

	Schooling			Earnings		
	Group Coefficients		p-value	Group Coefficients		p-value
Panel A: Top vs. Bottom Half Family SES						
	Low	*High*		*Low*	*High*	
Attention	-0.149***	-0.227***	0.420	-0.034***	-0.038***	0.539
	[0.052]	[0.074]		[0.012]	[0.013]	
Conduct	-0.046	-0.106*	0.324	0.038***	0.026**	0.123
	[0.039]	[0.062]		[0.009]	[0.011]	
Emotion	0.057	0.088	0.570	-0.018**	-0.030***	0.221
	[0.035]	[0.059]		[0.008]	[0.010]	
Peer	-0.013	0.042	0.422	-0.021**	-0.010	0.139
	[0.041]	[0.050]		[0.010]	[0.011]	
Cognition	0.622***	1.205***	0.000	0.113***	0.129***	0.069
	[0.049]	[0.054]		[0.010]	[0.011]	
N individuals	3,658	3,294		3,658	3,294	
N individual-years				11,850	11,601	
Panel B: College Degree vs. No Degree						
	None	*Degree*		*None*	*Degree*	
Attention	-0.156***	-0.387***	0.198	-0.033***	-0.057*	0.005
	[0.044]	[0.148]		[0.009]	[0.030]	
Conduct	-0.059	-0.102	0.450	0.033***	0.033	0.984
	[0.036]	[0.112]		[0.007]	[0.023]	
Emotion	0.064*	0.026	0.446	-0.025***	-0.021	0.483
	[0.034]	[0.092]		[0.007]	[0.020]	
Peer	0.011	0.026	0.879	-0.012*	-0.028	0.207
	[0.039]	[0.095]		[0.007]	[0.019]	
Cognition	0.819***	1.114***	0.001	0.126***	0.108***	0.030
	[0.037]	[0.105]		[0.008]	[0.021]	
N individuals	5,922	1,030		5,922	1,030	
N individual-years				19,743	3,708	
Backg. controls (apply to A & B)	X	X		X	X	

Notes: Data from BCS70. Each column reports measurement error–corrected estimates from regressions of log monthly earnings and years of schooling on standardized socio-emotional skills and cognition constructed from our dedicated measurement system. Estimates are shown separately by family socioeconomic background: Panel A: bottom vs. top half of the family-SES score distribution, and Panel B: parental education (None vs. Degree). All specifications control for gender, number of siblings, indicators for first-born, no father at birth, teenage mother, and region fixed effects. Earnings specifications additionally include gender-by-year and year fixed effects. Standard errors in square brackets are estimated from 250 bootstrap replications clustered at the individual level: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

F.2 Kitagawa-Oaxaca-Blinder Decomposition Framework

To quantify how much of the earnings gap between children from different socioeconomic backgrounds can be attributed to differences in childhood socio-emotional and cognitive skills, we employ the Kitagawa-Oaxaca-Blinder decomposition method (see e.g. Fortin et al., 2011).

Let Y_i represent the outcome of individual i (log earnings or years of schooling), and let $S_i \in \{H, L\}$ indicate whether individual i comes from a high-SES (H) or low-SES (L) family background. We estimate separate equations for each group as follows:

$$Y_i^H = X_i^H \beta^H + \varepsilon_i^H \quad \text{for } S_i = H \tag{7}$$

$$Y_i^L = X_i^L \beta^L + \varepsilon_i^L \quad \text{for } S_i = L \tag{8}$$

where X_i is a vector containing: childhood socio-emotional skills (attention, conduct, emotion, and peer problems); cognitive ability measured at age 10; and additional controls, including gender, birth order and region.

The observed earnings gap between high-SES and low-SES individuals is:

$$\Delta = \bar{Y}^H - \bar{Y}^L = \bar{X}^H \beta^H - \bar{X}^L \beta^L$$

We decompose this gap into an 'explained' part – due to differences in skill endowments – and an 'unexplained' part – due to differences in the returns to skills. We evaluate the difference in skill endowments using a coefficient vector β^* obtained from a pooled regression that includes group dummies for family SES. The coefficients from this pooled regression can be interpreted as the average returns that would prevail in the absence of structural barriers between groups. As shown in Table F.2, with the exception of attention problems for degree vs no degree, returns to socio-emotional skills do not differ significantly across SES groups, supporting the use of pooled coefficients as the reference.

$$\Delta = \underbrace{(\bar{X}^H - \bar{X}^L)\beta^*}_{\text{Explained}} + \underbrace{\bar{X}^L(\beta^L - \beta^*) + \bar{X}^H(\beta^* - \beta^H)}_{\text{Unexplained}} \tag{9}$$

$$= \Delta_E + \Delta_U \tag{10}$$

For each childhood skill k (where k ranges over: attention, conduct, emotion, peer and cognition), we further decompose the endowments effect as follows:

$$\Delta_{E,k} = (\bar{X}_k^H - \bar{X}_k^L)\beta_k^* \tag{11}$$

This allows us to quantify the contribution of each specific childhood skill to the explained portion of the SES schooling or earnings gap.

Estimation and Inference

As before, the coefficient estimates need to be corrected for measurement error. Similarly to our main analysis, we use the attenuation matrix derived from the covariance structure of the latent factors to correct the pooled vector of coefficients β^*. So, to calculate the explained component of the overall gap, we implement the following steps:

1. Estimate a pooled OLS regression with SES group dummies to obtain the vector of coefficients β^*;

2. Compute bias-corrected coefficients for β^* using the method described in Appendix A;

3. Calculate the contribution of each individual skill using sample means (Table F.1) and eq(11).

We do not explicitly show the unexplained component of the decomposition, as we see no evidence of systematic differences in the group specific coefficients, as shown in Table F.2. Standard errors of the coefficients are obtained by bootstrapping (250 replications), where we cluster observations at the individual level.