INSTITUTE OF ECONOMICS
Centre for Economic and Regional Studies

Közgazdaság- és Regionális Tudományi Kutatóközpont
KÖZGAZDASÁG-TUDOMÁNYI INTÉZET

KRTK-KTI WORKING PAPERS | KRTK-KTI MŰHELYTANULMÁNYOK

Intergenerational educational mobility – the role of non-cognitive skills

ANNA ADAMECZ-VÖLGYI - MORAG HENDERSON - NIKKI SHURE

CERS-IE WP – 2022/37

September 2022

https://kti.krtk.hu/wp-content/uploads/2021/09/CERSIEWP202137.pdf

Acknowledgments: The Nuffield Foundation is an endowed charitable trust that aims to improve social wellbeing in the widest sense. It funds research and innovation in education and social policy and also works to build capacity in education, science and social science research. The Nuffield Foundation has funded this project (grant number EDO/43570), but the views expressed are those of the authors and not necessarily those of the Foundation. More information is available at www.nuffieldfoundation.org. We thank the participants of the Hungarian Society of Economics (HSE), the Scottish Economic Society (SES) and the European Society for Population Economics (ESPE) annual conferences and János Kiss Hubert, John Jerrim and Jake Anders for helpful comments.

Disclosure statement: This research was supported by the Nuffield Foundation (grant number EDO/43570).

ABSTRACT

While it has been shown that university attendance is strongly predicted by parental education, we know very little about why some potential 'first in family' or first-generation students make it to university and others do not. This paper looks at the role of non-cognitive skills in the university participation of this disadvantaged group in England. We find that conditional on national, high-stakes exam scores and various measures of socioeconomic background, having higher levels of non-cognitive skills, specifically locus of control, academic self-concept, work ethic, and self-esteem, in adolescence is positively related to intergenerational educational mobility to university. Our results indicate that having higher non-cognitive skills helps potential first in family university students to compensate for their relative disadvantage, and they are especially crucial for boys. The most important channel of this relationship seems to be through educational attainment at the end of compulsory schooling.

JEL Classification Codes: I24, J24

Keywords: socioeconomic gaps, intergenerational educational mobility, higher education, non-cognitive skills

Anna Adamecz-Völgyi
Institute of Economics, Centre for Economic and Regional Studies (KRTK KRTI), Toth Kalman u. 4, 1097 Budapest
and
UCL Social Research Institute, University College London, 27 Woburn Square, London WC1H 0AA
e-mail: a.adamecz-volgyi@ucl.ac.uk

Morag Henderson
UCL Social Research Institute, University College London, 27 Woburn Square, London WC1H 0AA
e-mail: morag.henderson@ucl.ac.uk

Nikki Shure
Corresponding author.
UCL Social Research Institute, University College London, 27 Woburn Square, London WC1H 0AA
and
Institute of Labor Economics (IZA), Schaumburg-Lippe-Str. 5-9, D-53113 Bonn.
e-mail: nikki.shure@ucl.ac.uk

Generációk közötti oktatási mobilitás – a nem-kognitív készségek szerepe

ANNA ADAMECZ-VÖLGYI - MORAG HENDERSON - NIKKI SHURE

ÖSSZEFOGLALÓ

Számos kutatás megmutatta, hogy a felsőoktatási részvétel fontos előrejelzője a szülői iskolázottság: a diplomás szülők gyerekei nagyobb valószínűséggel mennek egyetemre, mint azok a fiatalok, akik családjukban elsőként tennék ezt. Keveset tudunk azonban arról, mi magyarázza a felsőoktatási részvételt a nemdiplomás szülők gyerekei között. Ez a tanulmány a nem-kognitív készségek szerepét vizsgálja az egyetemi részvételben. Angliai adatokkal dolgozunk, és azt találjuk, hogy azonos iskolai vizsgaeredmények és családi háttér mellett nagyobb valószínűséggel mennek egyetemre azok a potenciálisan elsőgenerációs fiatalok, akik serdülőkorukban magasabb nem-kognitív készségekkel (belső kontrollhely, akadémiai önkép, munkaetika és önértékelés) bírtak. Eredményeink azt mutatják, hogy ezek a készségek segítik a fiatalokat abban, hogy a diplomás szülők gyerekeihez képest mért hátrányos helyzetüket kompenzálják, és különösen fontosak a fiúk számára. A nem-kognitív készségek és az egyetemi részvétel közötti összefüggés legfontosabb csatornáját a közoktatást lezáró, 16 éves korban írt iskolai vizsgák jelentik, ezért érdemes az egyetemi részvétel ösztönzését szolgáló szakpolitikai intézkedéseket már ezt megelőzően megkezdeni.

JEL Classification Codes: I24, J24

Keywords: társadalmi-gazdasági háttérből eredő hátrányok, generációk közötti oktatási mobilitás, felsőoktatás, nem-kognitív készségek

1. Introduction

Higher education (HE) brings substantial benefits to its graduates and society. A range of evidence shows that university graduates are more likely to be employed, earn higher wages and have better health than their peers who did not obtain a university degree (Card 1999; Blundell, Dearden, and Sianesi 2005; Dickson 2013; Oreopoulos and Petronijevic 2013). University participation, however, is graded by socio-economic status (SES): those from more advantaged backgrounds are more likely to attend university and graduate (Blanden and Machin, 2004; Britton *et al.*, 2016; Walker and Zhu, 2018). Furthermore, the literature suggests that while returns to higher education might be heterogeneous by social background, there is no evidence that they would be smaller for those from lower SES background. On the contrary, Brand and Xie (2010) find that those who are the least likely to go to university might gain the most by earning a degree, and Dearden, McGranahan, and Sianesi (2004) show that low-SES men enjoy significantly higher wage returns to a college degree than high-SES men. In England Adamecz-Völgyi, Henderson, and Shure (2020b) document that wage returns to graduation are not lower among those who are the first in their family to go to university than among those whose parents are graduates. Thus, increasing the share of disadvantaged students in higher education is not only one way to improve equity, but also expected to be a fruitful social investment which increases efficiency.

While it is documented that young people from disadvantaged backgrounds are less likely to go to university despite the potential gains, there is less evidence regarding why some disadvantaged young people beat the odds and attend university, but others do not. In this paper, we try to explain why some individuals achieve intergenerational educational mobility. To do this, we look beyond traditional measures of academic attainment and socioeconomic status to measures of non-cognitive skills, specifically locus of control, academic self-concept, work ethic, and self-esteem. We look at a specific group of disadvantaged youth: those whose parents did not go to university. We call this group *the potential 'first in family' or potential FiF*. Within this group, we make a distinction between those who graduate and thus become *first in family or FiF* university graduates and the rest, who match their parents without a university degree. We exploit a cohort study from England, Next Steps, which has been linked to administrative data on test scores from national high-stakes school examinations. Importantly, the study also includes rich data on family background, adolescent non-cognitive skills, and educational achievements.

We focus on the potential FiF for two reasons. First, it is a surprisingly large group. Henderson, Shure, and Adamecz-Völgyi (2020) show that over 80 percent of a recent cohort born in 1989/1990 in England are potential FiF, i.e. neither of their (step)parents earned a university

degree. Using the same data as this paper, they find that 52 percent of young people whose parents had university degree went on to graduate by age 25, while among the potential FiF, this share was only 22 percent. Second, Adamecz-Völgyi, Henderson, and Shure (2020a) show that being a potential FiF is the most important social-background-related barrier to university participation. They document that while potential FiF students are a heterogeneous group, being a potential FiF decreases the probability of university participation and graduation even after controlling for detailed measures of family background, early educational attainment (national test scores measured at age 11), secondary school progression (national test scores measured at age 16), and other standard measures of disadvantage, for example, Free School Meals (FSM) eligibility. Furthermore, they show that potential FiF is the most important predictor of university participation compared to other commonly used measures of disadvantage in education.

Up until now, there has been very little, clear evidence about why some potential FiF students make it to university while others do not. Early educational attainment has been shown to be an important factor for HE participation (Gorard *et al.*, 2017; Adamecz-Völgyi, Henderson, and Shure, 2020a). Young people who do not have the grades will have a much lower probability of making it to university. There is a range of literature that shows that young people from disadvantaged backgrounds, including potential FiF, have lower average secondary school performance than their peers whose parents are university graduates (Henderson, Shure, and Adamecz-Völgyi 2020; Woessmann 2004), and educational attainment is a driver of intergenerational mobility (Blanden, Gregg, and Macmillan 2007; Chetty et al. 2020). A growing empirical literature documents the role of non-cognitive skills, over and above cognitive abilities on several domains such as educational attainment, labour market success and health (Kautz et al. 2014). This literature has benefitted from rich survey data, which allows economists to include measures of non-cognitive skills in the education production function. Blanden, Gregg, and Macmillan (2007) also include measures of non-cognitive skills as mechanisms in their calculation of the intergenerational elasticity of wages, concluding that non-cognitive skills impact educational outcomes for a cohort of people born in 1970 in the UK, but that they do not directly affect wages. Thus, this paper builds on existing literature by examining the role of non-cognitive skills in intergenerational educational mobility to university.

The term *non-cognitive skills* is sometimes seen as vague. In the economics literature, non-cognitive skills might encompass various things such as personality characteristics, motivation, attitudes, efforts, emotions, etc. (as opposed to cognitive skills that refer to innate intelligence (IQ) or learnt abilities captured through test scores in math, for example) (Almlund et al. 2011). A rich literature documents the importance of non-cognitive skills for life outcomes. A key conclusion

of the literature is that non-cognitive skills truly are "skills" and not "traits" – they can be developed and thus are potentially interesting for education policy. This is especially true if we consider that cognitive skills may be less malleable after a certain age while non-cognitive skills might stay malleable throughout adolescence and beyond (Kautz et al. 2014). Empirical evidence shows that certain policy interventions can influence non-cognitive skills (O'Mara et al. 2006) and thus potentially counteract the negative effects of low parental education on their accumulation. Looking at a cohort of university entrants at an Australian university, Schurer et al. (2020) find that non-cognitive skills, in particular, Conscientiousness, one of the Big Five personality traits, help first in family students to compensate for the academic penalties produced by social origin. However, there is a gap in the literature on the role of non-cognitive skills in university entry.

In this paper we focus on four key non-cognitive skills: locus of control, academic self-concept, work ethic, and self-esteem. Heckman, Stixrud, and Urzua (2006) find that high inner locus of control (when one believes that they themselves are responsible for their life outcomes as opposed to luck or faith) and self-esteem just as strongly influence schooling decisions as cognitive skills, while Jacob (2002) shows that some non-cognitive factors (effort put into learning, which may be related to work ethic) influence college enrolment even after controlling for education attainment in high school. Prada and Urzúa (2017) show that locus of control, self-esteem, and a measure of adolescent reckless behaviour increase the probability of four-year college attendance (as well as wages), after controlling for cognitive abilities. Academic self-concept, the belief in one's own academic ability, has been shown to be associated with an increase in educational attainment (e.g. Hansen and Henderson 2019; Valentine, DuBois, and Cooper 2004; Chowdry, Crawford, and Goodman 2011). Prevoo and ter Weel (2015) use a British cohort study, the BCS70, to investigate the association between conscientiousness and several outcomes. They find that conscientiousness is associated with higher levels of education, as well as fewer unhealthy behaviours, greater savings, and higher life satisfaction. Conscientiousness has been found to be closely related to work ethic (Mendolia and Walker 2014). Walker and Buchmueller (2020) use the same dataset as this paper and show that locus of control, work ethic, and self-esteem do not drive the graduate wage premium because they contribute equally to the wages of university graduates and non-graduates. Taken together this literature shows that the four skills we explore in this paper have predictive power in explaining university participation and other life outcomes, but it does not show how they contribute to intergenerational educational mobility and socioeconomic gaps in university participation.

Our contribution to the literature is fourfold. First, we quantify the gap in four adolescent non-cognitive skills (locus of control, academic self-concept, work ethic, and self-esteem) across four

groups of young people: (1) those who match their parents with no HE degrees, i.e. potential FiF who do not go to university; (2) the downwardly mobile group: young people who do not go to university despite their parents being graduates; (3) the upwardly mobile group: potential FiF who go to university and thus became actual FiF; (4) those who match their parents with HE degrees: young people who go to university and whose parents are graduates (see Section 3.1 for more information on these four groups). Interestingly, we find that after controlling for early educational attainment and family background, FiF university graduates possess the highest levels of adolescent non-cognitive skills in terms of locus of control, academic self-concept, and work ethic. These results indicate that the potential FiF can compensate for their relative disadvantage by having higher non-cognitive skills compared to young people whose parents are graduates. This would imply that the price of university entry in terms of human capital resources is higher for those who need to break the educational ceiling of their family than for those whose parents had already made this educational transition before them.

Our second contribution is that we look at the probability of HE participation among the potential FiF. We are interested in whether having higher non-cognitive skills explains why some potential FiF students go to university while others do not. We show that young people with lower external locus of control, and higher academic self-concept, work ethic and self-esteem in early adolescence are more likely to go to university, even after controlling for national exam scores at age 11 (Key Stage 2 or KS2), as a proxy for cognitive skills, and rich measures of individual and family background. Looking at the heterogeneity of these associations along the distribution of age 11 test scores reveals that external locus of control matters the most in the middle of the distribution, academic self-concept matters along the whole distribution, and work ethic is especially important in the upper-middle quintile. Having low self-esteem is only harmful to those at the top of the ability distribution.

The third contribution is to quantify the relative roles of non-cognitive skills and early educational attainment in the FiF gap in university participation. We decompose the gap in the probability of university participation between the potential FiF and children of graduate parents using the Kitagawa-Blinder-Oaxaca decomposition. We find that the contribution of academic self-concept to the difference in participation rates is substantially larger among boys than among girls. Thus, non-cognitive skills are especially important in explaining the socioeconomic university participation gap for boys.

The final contribution of the paper is to investigate a potential channel of the effects of childhood non-cognitive skills on university participation: test scores taken at the end of compulsory schooling, at age 16 (GCSEs), by adding these scores to the decomposition model.

We find that once the decomposition is extended to include age 16 test scores, they explain the largest share of the university participation gap. Lastly, as age 16 test scores might already be the consequence of adolescent non-cognitive skills, we decompose the age 16 test score gap between the potential FiF and children of graduate parents. We find that on average, adolescent academic self-concept explains more than two times as large a share of the gap as age 11 test scores among boys, while among girls, the contributions of non-cognitive skills and age 11 test scores are similar in magnitude. This implies that non-cognitive skills in adolescence shape not only the probability of attending university, but also the necessary prior attainment required to make this transition.

The rest of the paper unfolds as follows. Section 2 describes the data and Section 3 explains our empirical methods. Section 4 presents our results while Section 5 provides a discussion and conclusion.

2. Data

We use the Next Steps (formerly the Longitudinal Study of Young People in England, LSYPE) study which follows a cohort of young people born in 1989/1990 in England. Next Steps began in 2004 when the sample members were aged 13, with the most recent sweep of data collection at age 25 (University College London, UCL Institute of Education, Centre for Longitudinal Studies 2018). The data are linked with the National Pupil Database (NPD) which provides a census of pupils attending schools in England, allowing us to access their national school exam results.

Respondents of the Next Steps study were selected to be representative of young people in England using a stratified random sample of state and independent schools, with disproportionate sampling for deprived schools, i.e., those in the top quintile of schools in terms of the share of pupils eligible to FSM (Department for Education, 2011). Design weights were constructed to take care of the oversampling of deprived schools using inverse probability weighting (Department for Education, 2011). Some schools that were chosen to be in the sample decided not to participate. The first wave thus started with a 21,000-observation issued sample in 28 independent and 646 maintained schools with an average response rate of 74%, resulting in a 15,770-observation initial sample. Starting from Wave 1, attrition weights are estimated by stratum to take care of the initial school-level non-compliance as well as individual attrition from the study. All results that we present in this paper are estimated using the final weights available in the data.

Schools are the primary sampling units of Next Steps, then pupils within schools. The two-stage sampling design presents a possible clustering effect due to school-specific unobserved random shocks; therefore, we apply robust standard errors clustered by schools in all models (Abadie et al. 2017). In the first four waves both young people and their parents were interviewed,

and the information content of all variables on family background and parental education that we use in this paper was reported directly by the parents. From Wave 5, only young people were interviewed.

In terms of information on HE participation, we make use of the Next Steps age 25 data. The age 25 wave of Next Steps covers 7,707 young people, 36.7% of the initially drawn sample and 49% of the actual sample of the first wave. In Wave 4, an ethnic boost sample were added to the study, selected from the schools that were chosen at the beginning but did not participate in Wave 1 (Centre for Longitudinal Studies, 2018). As measures of non-cognitive skills are missing for these individuals, we exclude them from the sample (this means 138 individuals out of 7,707), along with 33 further individuals for whom information on parental education is missing.

2.1 Non-cognitive skills

We use non-cognitive measures from adolescence, collected in the first four waves of Next Steps, at ages between 13 and 16, to construct four indexes of non-cognitive skills: external locus of control, academic self-concept, work ethic and self-esteem. We use the earliest available data and decrease the share of missing values (that is between 0-5%) by replacing them from later waves (see Table A1 in Appendix A for details). Importantly, all these measures are collected before the individuals would have been applying to university and are described in more detail below.

External locus of control captures whether one believes that external circumstances, like luck or faith, are responsible for the outcomes of their life, and not they themselves (Rotter 1966). Having high external locus of control has been shown to be negatively associated with numerous educational, behavioural, labour market, and health outcomes (Mendolia and Walker 2014). Therefore, we expect higher levels of external locus of control to be negatively related with university participation. Following Schurer (2017) and Mendolia and Walker (2014), we conduct a principal component analysis (PCA) on the answers to six questions on locus of control from the first wave as listed in Appendix A. We use the first resulting factor as an index of external locus of control and standardise it to mean 0 and standard deviation 1. This allows us to interpret our results as the change in probability in going to university for a one standard deviation increase in locus of control. As the Cronbach's alpha is below 0.6 (0.44), we provide a robustness check using Confirmatory Factor Analysis (CFA) as detailed at the end of this subsection.

Academic self-concept is a student's general perception of their ability in school[1] (Gutman and Schoon 2016). Empirical evidence shows that self-concept of ability is malleable in school; interventions targeting the development of academic self-concept are effective on average (O'Mara et al. 2006). We expect higher values of academic self-concept to be positively related with university participation since a positive perception of ability in school should mean that individuals want to continue their education. We use questions on how individuals perceive their school achievement (Appendix A) to construct a standardised measure of academic self-concept using a PCA (Cronbach's alpha: 0.68). Out of the six questions, five come exclusively from the age 13 wave; while in the case of one question, 95.5% of the answers come from the first wave, 3.8% come from the second wave and 0.7% come from the third wave (Table A1 in Appendix A).

Work ethic is closely related to conscientiousness, the first of the Big Five personality inventory (Mendez and Zamarro 2016), which has been shown to positively influence many educational and other outcomes (Almlund et al. 2011). Therefore, we expect higher work ethic to be positively related to university participation since these individuals should have higher motivation and drive for applying to university. We create a standardised index of work ethic using a PCA (Cronbach's alpha: 0.56) on three questions that capture hard work and the importance of school success (Appendix A). Two questions come exclusively from the second wave: while for the third question, 96%, 3.6%, 0.4% of the answers come from the first, second and third waves, respectively (Table A1 in Appendix A).

Self-esteem captures one's perceptions on their own value (Coopersmith 1959). Following Mendolia and Walker (2014), we use two questions from the age 14 and age 16 waves that capture how useful/worthless one perceives themselves. In this case, as we only have two questions, we do not use PCA but define individuals having a low self-esteem if they put themselves in the lowest category at least once. We expect low self-esteem to be negatively related to university participation since individuals must have the confidence in their value to apply to university. In the econometric models, we use this as a binary measure, but on our graphs in Section 0, we make it visually more comparable to the other three indexes by standardisation. Thus, on the graphs in Section 0, higher values mean higher self-esteem.

Table 1 shows the descriptive statistics of non-cognitive skills (standardised across the total sample to mean 0 and SD 1) and the correlation matrix of the four indexes is shown in Table A2 in Appendix A. We estimate our main results on the complete case subsample of the data in terms

[1] Self-efficacy is a similar skill, but while self-concept focuses on how individuals feel about their past (or recent) achievements, self-efficacy measures expectations about their future performance (Gutman and Schoon 2016). In this paper we only look at self-concept.

of non-cognitive skills (4,755 individuals in the total sample and 3,775 individuals in the sample of potential FiF). Since PCA only produces a score for an individual if they gave an answer to all questions that belong to a measure, the first three non-cognitive skill measures have a high share of missing values. Thus, we provide a robustness check using Confirmatory Factor Analysis (CFA) estimated via Full Information Maximum Likelihood (*Structural Equation Modeling Reference Manual,* 2017). This method allows us to predict a score for all individuals who gave an answer to at least one question that belongs to a measure. The correlation between the PCA and CFA version of the measures is well above 0.9 (Table A2 in Appendix A). In Section 4, we provide results with PCA on the complete case sample while we replicate the main results with CFA on the largest possible sample in Appendix B.

2.2 Higher education participation

We look at the relationship between adolescent non-cognitive skills and the probability of higher education participation by age 25. Information on HE participation is supplied as a derived variable in Wave 8 (the most recent, age 25 wave of data) based on related information gathered from all waves and has no missing values. University participation is defined broadly as ever been to university, independent from the length of university attendance.

In the total sample, the weighted average of university participants is 40 percent (Table 1, Total sample), which is very similar to the cumulative ratio of the Higher Education Initial Participation Rate (HEIPR)[2] for this cohort, 40.4 percent, calculated by the Department for Education for English domiciled young people aged 17-25 (Department for Education 2017). Among the potential FiF, 33 percent, while among the children of graduate parents, 72 percent have ever been to university by age 25 (Table 1, Total sample).

Table 1: Sample descriptive statistics: means and proportions of variables of interest

	Total sample			Complete case sample			Difference between the total and the complete case sample	t-test p-value
	Obs	Mean	SD	Obs	Mean	SD		
All individuals								
External locus of control (continuous, standardized)	4,937	0	1	4,755	-0.02	0.99	0.02	0.26

[2] The HEIPR is applicable to students who live in England, enter HE for the first time, attend UK HE institutions and English, Welsh and Scottish further education colleges, and, stay in HE for at least six months (Department for Education 2017). It is an estimate of the likelihood of a young person participating in HE, in our case, by age 25, i.e. the sum of the likelihoods of HE entry at ages 17-25. This is estimated using the administrative data of the Higher Education Statistics Agency on the number university participants and the size of each cohort from the UK Census (Office for National Statistics). As the yearly publications of the Department for Education cover the data of those at age 17-25 in each year, we constructed the HEIPR rate of our particular cohort (those born in 1989-90) by using the appropriate year of observation for all ages (i.e., the data of the 2014/15 academic year for age 25, the data of the 2013/14 academic year for age 24, …, etc.).

Academic self-concept (continuous, standardized)	7,259	0	1	4,755	0.06	0.98	-0.06***	0.00
Work ethic (continuous, standardized)	6,538	0	1	4,755	0.02	0.99	-0.02	0.39
Low self-esteem (binary)	7,507	0	1	4,755	-0.01	1.01	0.01	0.76
HE participation (binary)	7,536	0.40	0.49	4,755	0.42	0.49	-0.02***	0.01
Potential FiF: young people whose parents do not have university degrees								
External locus of control (continuous, standardized)	3,906	0.05	1.01	3,753	0.03	1.00	0.02	0.29
Academic self-concept (continuous, standardized)	5,834	-0.06	0.99	3,753	0.00	0.96	-0.06***	0.00
Work ethic (continuous, standardized)	5,211	-0.01	1.00	3,753	0.00	0.99	-0.02	0.46
Low self-esteem (binary)	6,048	-0.01	1.01	3,753	-0.02	1.03	0.01	0.52
HE participation (binary)	6,073	0.33	0.47	3,753	0.35	0.48	-0.02**	0.05
Children of graduate parents								
External locus of control (continuous, standardized)	1,031	-0.23	0.93	1,002	-0.24	0.93	0.01	0.77
Academic self-concept (continuous, standardized)	1,425	0.31	1.01	1,002	0.37	1.00	-0.06	0.17
Work ethic (continuous, standardized)	1,327	0.06	0.98	1,002	0.08	0.98	-0.02	0.69
Low self-esteem (binary)	1,459	0.06	0.92	1,002	0.08	0.89	-0.02	0.52
HE participation (binary)	1,463	0.72	0.45	1,002	0.73	0.44	-0.01	0.51

Notes: *** p<0.01, ** p<0.05, * p<0.1 Weighted using Wave 8 weights.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016. http://doi.org/10.5255/UKDA-SN-5545-7

2.3 Control variables

In our main empirical models, we control for variables that are expected to affect university participation but could not have been affected by adolescent non-cognitive skills. This includes the following variables:

Demographics and family background: gender, age, number of siblings, fixed effects (FE) for the region of school at age 13, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, an area-level measure of income deprivation (IDACI), disability, whether or not their grandparent(s) attended university, FSM eligibility status, whether they went to private school, and highest qualifications of the mother and the father (only in models where parents are non-graduates). For all missing values for these variables, we include a missing flag.

We are interested in whether non-cognitive skills affect university participation above and beyond cognitive skills. We do not have a direct measure of cognitive skills, but we observe the results of national exam scores in school. Thus, we control for national exam scores measured at

the end of primary school at age 11, Key Stage 2 (KS2) scores, as a proxy for cognitive abilities. Note that this educational attainment measure was measured years before the non-cognitive measures. In some models, we also control for high-stake exam test scores measured at age 16, capped linear GCSE (Key Stage 4) scores[3], even though they might have been affected by adolescent non-cognitive skills and thus be *bad controls,* which might lead to biased estimates of the coefficients on non-cognitive skills (Angrist and Pischke 2008). Controlling for age 16 test scores might help to absorb any potential omitted variable bias, while as they are measured years after non-cognitive skills, they could also act as a channel through which non-cognitive skills are related with university participation. If we control for rich measures of social background, age 11 test scores and non-cognitive skills, there are probably few remaining omitted variables. Moreover, as non-cognitive skills, age 16 test scores and university participation are all positively correlated with each other, adding age 16 test scores to the models will attenuate the estimated coefficients on non-cognitive skills further. Thus, we interpret the estimated coefficients from these models as the lower bounds of the statistical relationship between non-cognitive skills and university participation. We use all test scores as categorical variables based on the quintiles of achievement and set a sixth category to capture their missing values.

3. Empirical methods

3.1 Graphical comparison

We start with a visual comparison where we plot the average of non-cognitive skills for each of the four groups mentioned earlier (Table 2).

Table 2: Comparison groups used in this paper

	Potential FiF (children of non-graduate parents)	Children of graduate parents
Young people who did not go to university	Group 1: those matching their parents with no HE (N=2,581)	Group 2: Downwardly mobile group (N=450)
Young people who went to university	Group 3: FiF (N=1,172)	Group 4: those matching their parents with HE (N=552)

The number of observations (N) refers to the complete case sample.
Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016. http://doi.org/10.5255/UKDA-SN-5545-7.

[3] This linear measure is created by assigning values to the grades, for example, Grade G is worth 16 points. Each grade improvement thereafter, for example, from G to F, C to B, or A to A*, is equivalent to an additional six points. The capped linear score takes the best eight GCSE subjects scores. This measure takes into account the fact that students may take a different number of GCSEs (and resits) and enables better comparability than the total GCSE score.

We plot the raw as well as the conditional means of non-cognitive skills. We construct the conditional means by estimating linear regression models with one of the standardised indexes of non-cognitive skills at a time on the left hand-side of the equation, and the following control variables on the right hand-side: birth weight; whether the individual was born prematurely; age in months at the time of the survey (age 25 wave); ethnicity; fixed effects (FE) for the region of school at age 13; mother's and father's age and social class at age 13; the highest qualification of the mother and the father (only if parents are non-graduates); whether the grandparents of individuals have university degree; the number of siblings; an area-level measure of income deprivation (IDACI); disability; whether the individual has ever been in care; FSM eligibility; living in a single household; and high-stakes tests results taken at age 11 (KS2 total score quintiles). Then, we predict the outcome variables from the models and plot the residuals for each of the four groups.

3.2 Modelling intergenerational educational mobility

We are interested in whether certain non-cognitive skills decrease or increase the probability of educational mobility among potential FiF individuals. In an ideal world, we would want to compare the educational outcomes of young people in a random experiment in which *low* vs. *high* levels of non-cognitive skills have been randomly allocated to the individuals. As we cannot exploit such a random experiment, we face a selection problem. Those with *low* vs. *high* levels of non-cognitive skills might genuinely differ from each other on many other observed and unobserved domains, and this difference might be related to their educational outcomes. While we cannot fully solve the problem of unobserved selection, we make use of a rich dataset that allows us to control for several observed sources of selection (e.g. family background, educational attainment).

We estimate linear probability models as:

$$university\_attendance_i = \propto + \beta * non\_cognitive\_measures_i + \gamma * X_i + \varepsilon_i \quad (1)$$

where

$university\_attendance_i$ is a binary variable capturing whether individual i ever attended university by age 25/26;

$non\_cognitive\_measure_i$ is one of (or a vector of) the non-cognitive measures of individual i;

X_i is a vector of individual characteristics for individual i;

ε_i is an individual-level error term, robust and clustered by sampling schools.

We first estimate models that include one non-cognitive measure at a time (Model 1-4) and the control variables listed in Subsection 2.3.[4] Then, we re-estimate the model including all four non-cognitive measures in the model at once. Lastly, although it might be *bad control* as discussed earlier, we test whether non-cognitive skills still matter conditional on age 16 test scores.

Then, following Kalil and Khalid (2010), we re-estimate the main model by age 11 test score quintiles to investigates whether adolescent non-cognitive skills are more or less important to those having lower versus higher early educational attainment. Recall that all non-cognitive measures are measured from age 13 onwards, so after age 11 exams. Since men and women have different non-cognitive skills and associated mechanisms (Almlund et al. 2011), we replicate these results separately for boys and girls in Subsection 4.3.

3.3 Kitagawa-Blinder-Oaxaca decomposition

In the next step, we decompose the difference in the probability of university attendance between potential FiF and children of graduate parents using the Kitagawa-Blinder-Oaxaca decomposition (Kitagawa 1955; Oaxaca 1973; Blinder 1973). The method decomposes the difference in an outcome variable across two groups to an *explained part* that is due to the different distributions of explanatory variables in the two groups (*endowments*), and to an *unexplained* part that is not attributable to the distributions of the explanatory variables (rather it captures the *returns* to those characteristics). We apply common coefficients estimated from a pooled regression (Neumark 1988),[5] thus, the estimated coefficient of the unexplained gap is identical to the coefficient of potential FiF in a regression model that pools together the data of the two groups and controls for potential FiF as well as the same control variables (Słoczyński 2020). In other words, the unexplained gap is the gap that remains after including all control variables. The value added of this method compared to a regression is that it shows how large is the relative contribution of each non-cognitive skill to the raw gap in one step.

3.4 Testing a potential channel: compulsory school leaving exams

Lastly, we extend the Kitagawa-Blinder-Oaxaca decomposition with age 16 test scores from national, compulsory school leaving exams (GCSEs). This exercise is not straightforward, because as discussed earlier, test scores at age 16 could already be the consequence of having different non-

[4] The hierarchical structure of the data raises the question of applying fixed effects (FE) (or hierarchical) models. However, as in our main estimation sample the number of observations is 3,753 and the sample was taken from 674 schools, the average number of observations per school is lower than the number of explanatory variables in the model. Thus, we decided not to present the results from FE models (which were similar in terms of magnitude). This similarity reflects the fact that we control for socio-economic background extensively and it is unlikely that there is systematic selection to schools based on non-cognitive skills.

[5] We also experimented with the binary-dependent-variable-extension of the decomposition as in Fairlie (2005) and it yielded similar results.

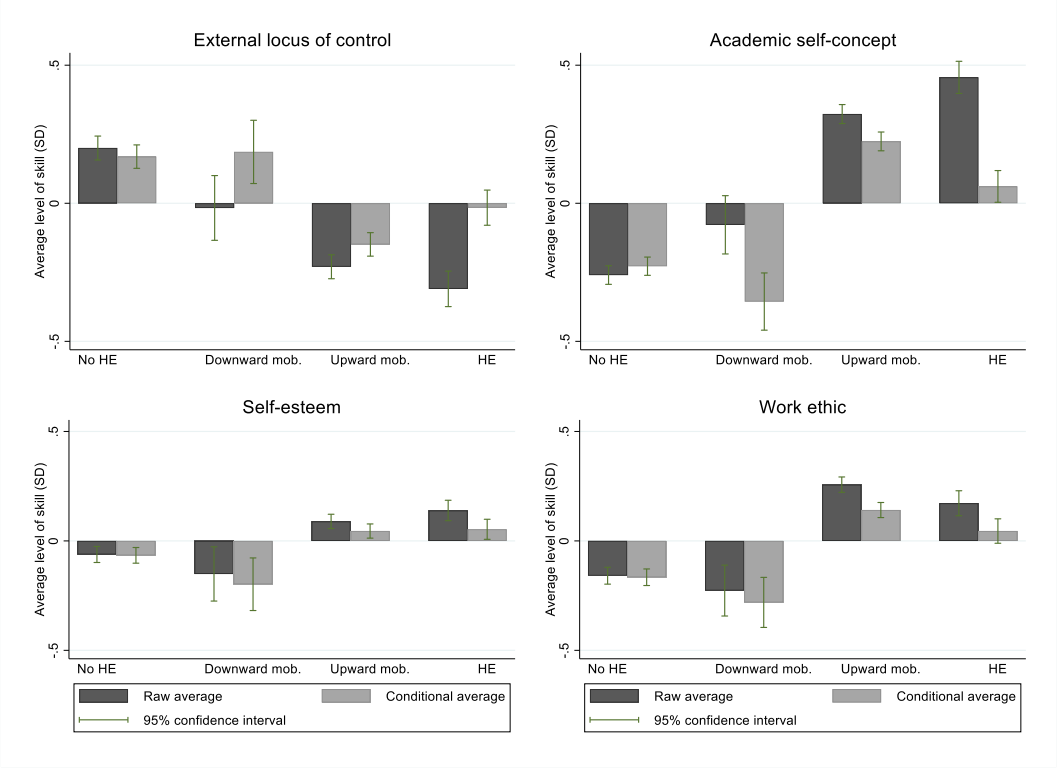
cognitive skills at age 13. Note that the educational attainment measure that we have used in our main models was measured at age 11, years before the non-cognitive measures. We add age 16 test scores to the models to investigate whether they could be a potential channel of the effects of non-cognitive skills on university participation. Since we find that they are, we replicate the same decomposition exercise placing age 16 test scores on the left hand-side to show that non-cognitive skills are an important source of the variation in test scores between the potential FiF and children of graduate parents.

4. Results

4.1 Graphical comparison

First, we compare the unconditional and conditional means of non-cognitive skills across the four groups defined in Table 2. Figure 1 shows that those who went to university, the upwardly mobile group of FiF university graduates (Group 3) and the group whose parents are graduates (Group 4), had substantially higher levels of non-cognitive skills in adolescence even after controlling for individual characteristics and prior educational attainment, than those who did not go to university (Groups 1 and 2). Interestingly, in three out of the four domains, FiF young people (Group 3) had significantly higher non-cognitive skills even than those university entrants whose parents are graduates (Group 4). The FiF individuals have the lowest external locus of control, the highest academic self-concept and work ethic, while in terms of self-esteem, which is standardised to be visually comparable to the other three measures, they are similar to the group whose parents are graduates (Group 4).

Figure 1: Non-cognitive skills, raw and conditional means



Notes: N = 4,937, 6,538, 7,259 and 7,507, respectively. Weighted using Wave 8 weights. Conditional means control for gender, age, number of siblings, region, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, income deprivation, disability, whether their grandparent attended university, Free School Mean (FSM) status, private school, the highest qualification of the mother and father, and Key Stage 2 quintile of achievement.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016: Secure Access. DOI: 10.5255/UKDA-SN-7104-4

When we look at the potential FiF group, FiF young people (Group 3) had about 0.3 standard deviations (SD) lower level of external locus of control, more than 0.4 SD higher level of academic self-concept, 0.3 SD higher level of work ethic and roughly 0.1 SD higher level of self-esteem in adolescent than those matching their parents with not going to university (Group 1), even after controlling for individual characteristics and prior educational attainment. Thus, this suggests that non-cognitive skills are important factors for intergenerational educational mobility.

4.2 Modelling intergenerational educational mobility

We now turn our attention to how the aforementioned non-cognitive skills predict university attendance for the potential FiF (Groups 1 and 3). As outlined in Section 3.2 we estimate a series of linear probability models where the outcome is a binary variable for attending university. The sample includes potential FiF individuals only, as we want to understand how non-cognitive skills facilitate intergenerational educational mobility for those whose parents are non-graduates.

The results of this analysis are presented in Table 3. In Models 1-4, we regress the binary outcome variable on each of the non-cognitive measures one at time while controlling for the background characteristics previously outlined. In Model 5, we enter all four non-cognitive measures at the same time and in Model 6 we additionally control for age 16 exam score quintiles. To address concerns about missing data, we replicate the main model, Model 5, using CFA instead of PCA to create the non-cognitive measures in Table B1 in Appendix B; the results are very similar. Table 3 shows that non-cognitive skills play an important role in facilitating intergenerational educational mobility. All non-cognitive skills are statistically significant predictors of university attendance on their own. In a joint model, when all measures are included at the same time, their magnitudes decrease, but they all remain statistically significant. One standard deviation higher external locus of control decreases the probability of university participation by 3.8 percentage points, while one standard deviation higher work ethic and academic self-concept increases it by 2.5 and 8.9 percentage points, respectively. Having low self-esteem is associated with 5.4 percentage points lower likelihood of university participation. These relationships are over and above prior attainment (age 11 test scores) and a range of socio-demographic characteristics, all measured prior to these non-cognitive traits. Even in Model 6, where we introduce age 16 exam performance, which is expected to bias the estimated coefficients on non-cognitive scores downwards in terms of magnitude, the statistical significance of work ethic and academic self-concept remain. One standard deviation higher work ethic is associated with 1.7 percentage point higher probability of university participation while one standard deviation higher academic self-concept is associated with 2.8 percentage point higher probability of university participation.

Table 3: The relationship between non-cognitive skills and the probability of university participation among the potential FiF

	(1)	(2)	(3)	(4)	(5)	(6)
External locus of control (continuous, standardized)	-0.071***				-0.038***	-0.003
	(0.008)				(0.008)	(0.008)
Work ethic (continuous, standardized)		0.071***			0.025***	0.017**
		(0.008)			(0.008)	(0.008)
Academic self-concept (continuous, standardized)			0.110***		0.089***	0.028***
			(0.009)		(0.009)	(0.008)
Low self-esteem (binary)				-0.099***	-0.054**	-0.033
				(0.023)	(0.023)	(0.022)
Constant	1.308**	1.389**	1.246**	1.436**	1.208**	1.152**
	(0.556)	(0.566)	(0.556)	(0.572)	(0.539)	(0.501)

Observations	3,753	3,753	3,753	3,753	3,753	3,753

Notes: Sample of potential first in family individuals (i.e. neither parent has a university degree). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Weighted using Wave 8 weights. All models control for gender, age, number of siblings, region, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, income deprivation, disability, whether their grandparent attended university, Free School Mean (FSM) status, private school, the highest qualification of the mother and father, and Key Stage 2 quintile of achievement. Column 6 additionally includes GCSE total points quintile. Model 5 using CFA instead of PCA to create the indexes of non-cognitive skills is reported in Table B1 in Appendix B.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016: Secure Access. DOI: 10.5255/UKDA-SN-7104-4

We explore whether the relationship between non-cognitive skills and university participation is heterogeneous along the distribution of prior educational attainment, test scores measured at age 11, in Table 4 (and in Table B2 in Appendix B using CFA). The association between external locus of control and university participation is the strongest in the middle of the distribution, academic self-concept matters along the whole distribution, and work ethic is especially important in the upper-middle quintile. Having low self-esteem is only harmful to those at the top of the ability distribution.

Table 4: The relationship between non-cognitive skills and the probability of university participation along the quintiles of age 11 test scores (KS2) among the potential FiF

	(1) Lowest quintile	(2) Lower-middle quintile	(3) Middle quintile	(4) Upper-middle quintile	(5) Highest quintile
External locus of control (continuous, standardized)	-0.027	-0.029	-0.065***	-0.029	-0.042
	(0.017)	(0.019)	(0.019)	(0.019)	(0.027)
Work ethic (continuous, standardized)	0.012	-0.002	0.015	0.059***	0.016
	(0.020)	(0.021)	(0.018)	(0.019)	(0.026)
Academic self-concept (continuous, standardized)	0.068***	0.110***	0.098***	0.103***	0.057**
	(0.021)	(0.021)	(0.021)	(0.022)	(0.026)
Low self-esteem (binary)	-0.033	-0.044	0.037	-0.057	-0.178***
	(0.047)	(0.054)	(0.051)	(0.055)	(0.058)
Constant	2.708**	-0.479	0.098	0.694	-0.231
	(1.156)	(1.315)	(1.264)	(1.310)	(1.460)
Observations	734	711	734	673	611

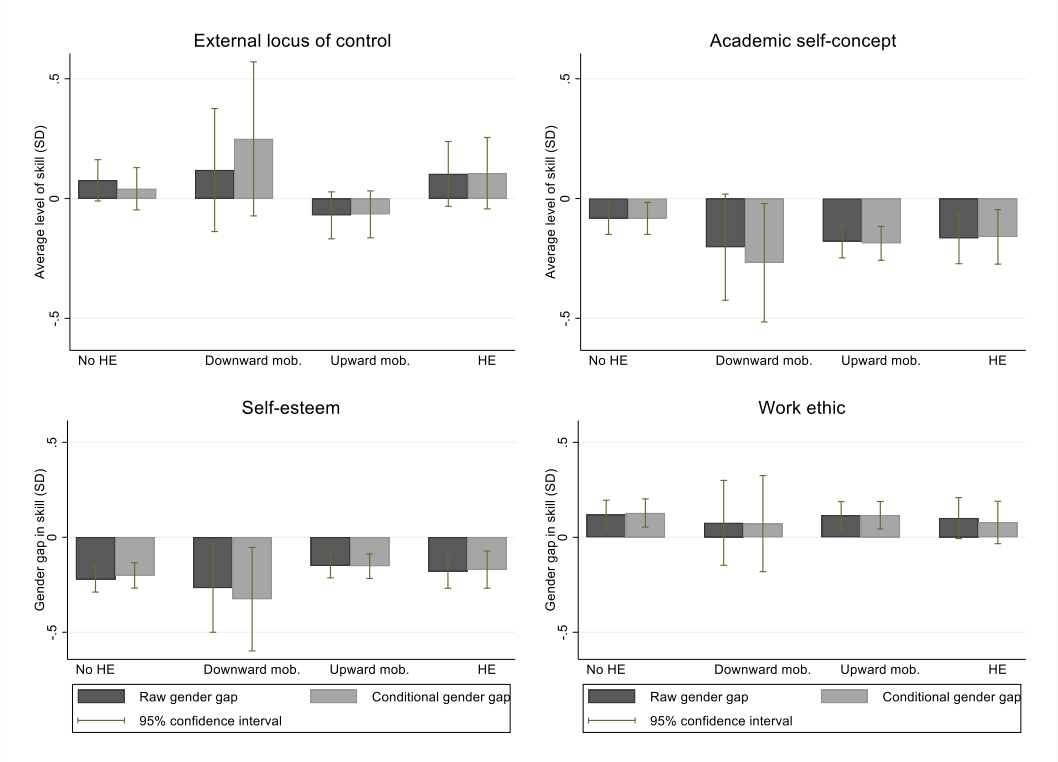
Notes: Sample of potential first in family individuals (i.e. neither parent has a university degree). Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Weighted using Wave 8 weights. All models control for gender, age, number of siblings, region, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, income deprivation, disability, whether their grandparent attended university, Free School Mean (FSM) status, private school, and the highest qualification of the mother and father. The same estimates using CFA instead of PCA to create the indexes of non-cognitive skills are reported in Table B2 in Appendix B.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016: Secure Access. DOI: 10.5255/UKDA-SN-7104-4

4.3 Heterogeneous effects by gender

In this section we look at the effects of non-cognitive skills separately by gender. First, we plot the gender gap in each non-cognitive skill for the four groups as specified in Table 1. Similarly to Figure 1, Figure 2 also shows the raw as well as the conditional gender gap (that is, the coefficient on female after controlling for age, number of siblings, region of residence during secondary school, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, income deprivation, disability, whether their grandparent attended university, Free School Mean (FSM) status, private school, the highest qualification of the mother and father, and Key Stage 2 quintile of achievement). The gender gap is insignificant in external locus of control in all groups. Girls have higher work ethic in three out of the four groups, and lower academic self-concept and lower self-esteem in all groups.

Figure 2: The gender gap in non-cognitive skills



Notes: Positive gaps indicate that women have higher level of a skill. Total N = 4,937, 6,538, 7,259 and 7,507, respectively. Weighted using Wave 8 weights. Conditional gender wage gap models control for age, number of siblings, region, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, income deprivation, disability, whether their grandparent attended university, Free School Mean (FSM) status, private school, the highest school qualification of the mother and father, and their Key Stage 2 quintile of achievement.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016: Secure Access. DOI: 10.5255/UKDA-SN-7104-4

Table 5 reproduces the models of Table 3, adding the interaction terms of female and the non-cognitive skills. These results show that having high external locus of control is more detrimental

for girls, while having high academic self-concept is somewhat more important for boys (Column 5). These results prevail even after controlling for GCSE scores as well (Column 6). Again, to address missing data concerns, we replicate the main model, Model 5, using CFA instead of PCA to create the non-cognitive measures in Table B3 in Appendix B.

Table 5: The relationship between non-cognitive skills and university attendance, by gender*non-cognitive skills interactions (sample of potential FiF)

	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.076***	0.064***	0.086***	0.083***	0.087***	0.036**
	(0.016)	(0.016)	(0.016)	(0.018)	(0.017)	(0.016)
External locus of control (continuous, standardized)	-0.055***				-0.022*	0.010
	(0.011)				(0.012)	(0.012)
Female* External locus of control	-0.035**				-0.037**	-0.028*
	(0.016)				(0.017)	(0.016)
Work ethic (continuous, standardized)		0.072***			0.022*	0.013
		(0.012)			(0.013)	(0.012)
Female* Work ethic		-0.003			0.003	0.008
		(0.015)			(0.017)	(0.015)
Academic self-concept (continuous, standardized)			0.134***		0.121***	0.051***
			(0.012)		(0.014)	(0.012)
Female*Academic self-concept			-0.046***		-0.062***	-0.043***
			(0.016)		(0.018)	(0.015)
Low self-esteem (binary)				-0.089**	-0.034	-0.014
				(0.037)	(0.037)	(0.034)
Female* Low self-esteem				-0.016	-0.026	-0.026
				(0.044)	(0.044)	(0.041)
Constant	1.301**	1.387**	1.259**	1.435**	1.217**	1.159**
	(0.556)	(0.566)	(0.557)	(0.572)	(0.540)	(0.502)
Observations	3,753	3,753	3,753	3,753	3,753	3,753

Notes: Sample of potential first in family individuals (i.e. neither parent has a university degree). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Weighted using Wave 8 weights. All models control for age, number of siblings, region, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, income deprivation, disability, whether their grandparent attended university, Free School Mean (FSM) status, private school, the highest qualification of the mother and father, and Key Stage 2 quintile of achievement. Column 6 additionally includes GCSE total points quintile. Model 5 using CFA instead of PCA to create the indexes of non-cognitive skills is reported in Table B3 in Appendix B.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016: Secure Access. DOI: 10.5255/UKDA-SN-7104-4

Looking at these interaction terms by KS2 quintiles (Table 6) reveals that external locus of control is only important for girls at the bottom of the skill distribution, while having low self-esteem is equally problematic for boys and girls at the top of the skill distribution. Furthermore, the coefficient on the female variable suggests that girls are increasingly more likely to go to university than boys moving up along the skill distribution. While in the lowest quintile, potential FiF girls are not significantly more likely to go to university than potential FiF boys (0.056), in the highest quintile, they are 16 percentage point more likely to attend. This pattern is similar using the CFA method as well (Table B4 in Appendix B).

Table 6: The relationship between non-cognitive skills and university attendance by KS2 quintile, gender interactions (sample of potential FiF)

	(1)	(2)	(3)	(4)	(5)
	Lowest quintile	Lower-middle quintile	Middle quintile	Upper-middle quintile	Highest quintile
Female	0.056	0.066*	0.078**	0.127***	0.160***
	(0.037)	(0.037)	(0.038)	(0.040)	(0.043)
External locus of control (continuous, standardized)	0.005	0.004	-0.078***	0.002	-0.047
	(0.021)	(0.027)	(0.026)	(0.029)	(0.039)
Female* External locus of control	-0.064**	-0.069*	0.019	-0.055	0.006
	(0.032)	(0.037)	(0.036)	(0.043)	(0.051)
Work ethic (continuous, standardized)	-0.008	0.000	0.004	0.049*	0.020
	(0.026)	(0.032)	(0.024)	(0.028)	(0.040)
Female* Work ethic	0.029	-0.003	0.012	0.016	-0.006
	(0.034)	(0.042)	(0.038)	(0.038)	(0.047)
Academic self-concept (continuous, standardized)	0.109***	0.154***	0.157***	0.153***	0.079**
	(0.029)	(0.032)	(0.032)	(0.033)	(0.034)
Female*Academic self-concept	-0.074*	-0.081**	-0.106***	-0.089**	-0.044
	(0.039)	(0.040)	(0.038)	(0.045)	(0.051)
Low self-esteem (binary)	0.018	-0.014	0.025	0.009	-0.162*
	(0.076)	(0.071)	(0.076)	(0.121)	(0.094)
Female* Low self-esteem	-0.068	-0.042	0.025	-0.094	-0.024
	(0.093)	(0.099)	(0.099)	(0.144)	(0.121)
Constant	2.648**	-0.648	0.726	0.823	-0.309
	(1.137)	(1.333)	(1.282)	(1.315)	(1.465)
Observations	734	711	734	673	611

Notes: Sample of potential first in family individuals (i.e. neither parent has a university degree). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Weighted using Wave 8 weights. All models control for gender, age, number of siblings, region, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, income deprivation, disability, whether their grandparent attended university, Free School Mean (FSM) status, private school, and the highest qualification of the mother and father. The same estimates using CFA instead of PCA to create the indexes of non-cognitive skills are reported in Table B4 in Appendix B.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016: Secure Access. DOI: 10.5255/UKDA-SN-7104-

4.4 Kitagawa-Blinder-Oaxaca decomposition

In this subsection we look at the relative contributions of non-cognitive skills to the gap in university participation between potential FiF and non-FiF young people. In all decompositions presented below, we control for the same variables as in Models 1-5 in Table 3. These are grouped into categories for the purposes of the decomposition according to the following logic:

- Region: a set of dummies for the region of residence at age 13;
- Birth: a set of dummies with birth weight category, a dummy variable for a premature birth, ethnicity dummies, and the age of the mother at birth;

- Family: the number of siblings, single parent household dummy, and a dummy for whether or not the individual's grandparents had a university degree;
- Disadvantage: a dummy for FSM eligibility, a dummy for being a care leaver, a dummy for having a disability, and dummies for household income deprivation;
- Parents: mother's and father's social class and age;
- Key Stage 2: quintiles of age 11 test scores;
- GCSE: quintiles of age 16 test scores.

Decomposing the gap in the probability of university participation between children of non-graduate and graduate parents reveals that endowments explain about 60% of the difference (0.187 out of 0.300, see Table 7). Out of the variable groups, differences in parental background (that is, beyond whether parents went to university) explain the largest share of the difference, 27%. The contribution of high-stakes test scores measured at age 11 explain 7% of the difference, while the contribution of academic self-concept is 10% and the contribution of external locus of control is 3%. The contribution of the other two skills, low self-esteem and work ethic is close to zero. Interestingly, academic self-concept seems to be more important for boy's university participation (13% vs. 7%) while external locus of control is somewhat more important for girl's (4% vs. 2%).

Table 7: Kitagawa-Blinder-Oaxaca decomposition of university attendance among the potential FiF and children of graduate parents

	Total sample	Boys	Girls
Group 1: children of graduate parents	0.773***	0.758***	0.789***
	(0.013)	(0.019)	(0.018)
Group 2: potential FiF	0.473***	0.430***	0.507***
	(0.008)	(0.012)	(0.011)
Difference between the two groups	0.300***	0.327***	0.282***
	(0.016)	(0.023)	(0.021)
Explained part	0.187***	0.194***	0.178***
	(0.012)	(0.019)	(0.017)
Unexplained part	0.114***	0.134***	0.105***
	(0.018)	(0.026)	(0.024)
Explained difference			
External locus of control	0.009***	0.005	0.010***
(continuous, standardized)	(0.002)	(0.003)	(0.003)
Work ethic	0.001	0.002	0.001
(continuous, standardized)	(0.001)	(0.001)	(0.001)
Academic self-concept	0.031***	0.043***	0.020***
(continuous, standardized)	(0.004)	(0.007)	(0.004)
Low self-esteem	0.001	0.001	0.001
(binary)	(0.001)	(0.001)	(0.001)
KS2	0.020***	0.019***	0.021***
	(0.005)	(0.006)	(0.007)
Family	0.024***	0.020***	0.027***
	(0.005)	(0.008)	(0.007)
Region	0.001	0.004	-0.002
	(0.002)	(0.003)	(0.002)

Disadvantage	0.016***	0.014**	0.015**
	(0.005)	(0.006)	(0.007)
Birth	-0.019***	-0.015**	-0.017***
	(0.005)	(0.006)	(0.006)
Parents	0.082***	0.087***	0.075***
	(0.011)	(0.017)	(0.015)
Independent school	0.020***	0.014**	0.027***
	(0.005)	(0.006)	(0.007)
Observations	4,755	2,150	2,605

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Weighted using Wave 8 weights. All models control for gender, age, number of siblings, region, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, income deprivation, disability, whether their grandparent attended university, Free School Mean (FSM) status, private school, and Key Stage 2 quintile of achievement. 'Group 1' refers to individuals whose parents have a university degree and 'Group 2' refers to those individuals whose parents do not have a university degree.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016: Secure Access. DOI: 10.5255/UKDA-SN-7104-4

4.5 Potential channel: compulsory school leaving exams at age 16

In this subsection we extend the decomposition exercise with compulsory school leaving exams taken at age 16. As mentioned before, these could already be affected by adolescent non-cognitive skills, hence we have not included them in our main model in the previous subsection. As expected, age 16 test scores make a substantial contribution to the difference in university participation of the potential FiF and children of graduate parents (Table 8). On average, age 16 test scores explain about 49 percent of the university participation gap. Conditional on age 16 test scores, the contribution of academic self-esteem is three percent while the contribution of age 11 test scores is reduced to just two percent. When both age 11 and age 16 test scores are included in a model, we could interpret the earlier measure as a *proxy for cognitive abilities* or *baseline educational attainment*, while age 16 test scores capture everything that happened to the educational attainment of a young person between age 11 and age 16; we could call this as a *measure of educational progression*. These results suggest that for university participation, age 16 educational attainment is the most important factor.

As we find that the contribution of age 16 test scores is the most important in the university participation gap, we turn to decomposing these using the same method. This allows us to determine how important our measures of non-cognitive skills are for this crucial intermediate outcome. We find that on average, potential FiF have 0.551 SD lower age 16 test scores than children of graduate parents (Table 8). In terms of education effect sizes, this is a substantial difference (Kraft 2020). Most of this gap, 27 percent, is explained by parental background factors (parents' social class and age). The contribution of academic self-concept is 15 percent, the contribution of external locus of control is six percent, while that of age 11 test scores is eight percent. Thus, the contribution of academic self-concept to age16 test scores is even a bit higher

than the contribution of early educational attainment. Looking separately by gender, the contribution of academic self-concept to age 16 test scores is higher for boys than girls (18 vs. 2%).

Table 8: Kitagawa-Blinder-Oaxaca decomposition of university attendance among the potential FiF and children of graduate parents: the role of age 16 test (GCSE) scores

	University participation			GCSE's (standardized)		
	Total	Boys	Girls	Total	Boys	Girls
Group 1: children of graduate parents	0.773***	0.758***	0.789***	0.795***	0.754***	0.835***
	(0.013)	(0.019)	(0.018)	(0.019)	(0.026)	(0.028)
Group 2: potential FiF	0.473***	0.430***	0.507***	0.243***	0.142***	0.324***
	(0.008)	(0.012)	(0.011)	(0.013)	(0.021)	(0.017)
Difference between the two groups	0.300***	0.327***	0.282***	0.551***	0.613***	0.511***
	(0.016)	(0.023)	(0.021)	(0.023)	(0.033)	(0.033)
Explained part	0.239***	0.255***	0.224***	0.366***	0.398***	0.341***
	(0.013)	(0.019)	(0.018)	(0.021)	(0.034)	(0.028)
Unexplained part	0.062***	0.072***	0.058***	0.186***	0.214***	0.170***
	(0.017)	(0.025)	(0.022)	(0.022)	(0.033)	(0.031)
	Explained difference					
External locus of control	0.000	-0.004	0.002	0.032***	0.042***	0.023***
(continuous, standardized)	(0.002)	(0.003)	(0.002)	(0.006)	(0.009)	(0.007)
Work ethic	0.001	0.001	0.001	0.001	0.002	0.000
(continuous, standardized)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Academic self-concept	0.009***	0.017***	0.002	0.083***	0.108***	0.061***
(continuous, standardized)	(0.003)	(0.005)	(0.003)	(0.009)	(0.015)	(0.011)
Low self-esteem	0.001	0.000	0.001	0.001	0.001	0.001
(binary)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
KS2	0.007	0.006	0.006	0.043***	0.045***	0.042***
	(0.004)	(0.006)	(0.006)	(0.006)	(0.009)	(0.009)
GCSE	0.148***	0.161***	0.139***			
	(0.008)	(0.013)	(0.011)			
Family	0.016***	0.014*	0.016**	0.039***	0.029***	0.046***
	(0.005)	(0.007)	(0.007)	(0.007)	(0.011)	(0.010)
Region	-0.001	0.002	-0.002	0.004	0.007	0.001
	(0.001)	(0.003)	(0.002)	(0.002)	(0.005)	(0.003)
Disadvantage	0.002	0.001	-0.001	0.013	0.013	0.011
	(0.005)	(0.007)	(0.008)	(0.010)	(0.014)	(0.014)
Birth	-0.013***	-0.014**	-0.011**	-0.023***	-0.009	-0.016**
	(0.004)	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
Parents	0.039***	0.044***	0.034**	0.149***	0.148***	0.140***
	(0.010)	(0.016)	(0.014)	(0.015)	(0.024)	(0.021)
Independent school	0.031***	0.026***	0.037***	0.023***	0.012	0.032**
	(0.006)	(0.008)	(0.008)	(0.008)	(0.010)	(0.013)
Observations	4,755	2,150	2,605	4,420	2,005	2,415

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Weighted using Wave 8 weights. All models control for gender, age, number of siblings, region, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, income deprivation, disability, whether their grandparent attended university, Free School Mean (FSM) status, private school, and their Key Stage 2 and GCSE quintile of achievement. 'Group 1' refers to individuals whose parents have a university degree and 'Group 2' refers to those individuals whose parents do not have a university degree.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016: Secure Access. DOI: 10.5255/UKDA-SN-7104-4

5. Discussion

This paper looked at the role of non-cognitive skills in facilitating intergenerational educational mobility. We have shown that comparing university participants, first in family young people

possess lower external locus of control, higher work ethic, and higher academic self-concept in early adolescence than those whose parents are graduates. Thus, we document that conditional on early education attainment, having higher non-cognitive skills helps potential first in family university students to compensate for their relative disadvantage and beat the odds in terms of going to university. Our results are in line with Schurer et al. (2020) who find that non-cognitive skills are important for compensating social disadvantage of first in family university students.

Looking at potential FiF young people specifically, we find that non-cognitive skills matter for university participation above and beyond early educational attainment and various measures of individual characteristics and social background. While the coefficients on these variables reduce once we condition on prior attainment and background characteristics, they remain statistically significant predictors of university participation, highlighting the important role for non-cognitive skills in facilitating intergenerational educational mobility. This remains true when all measures are included in one model.

When we turn our attention to heterogeneous effects, we find that while work ethic and academic self-concept seem to be important for everybody, and especially crucial for boys, low self-esteem harms only those at the top of the ability distribution, boys and girls alike. This points to a need for a special focus on high-achieving, but underconfident disadvantaged young people. Overall, our results indicate that non-cognitive and cognitive skills (as proxied by prior attainment) are not substitutes for each other, but rather complement each other: having high educational attainment is necessary but not sufficient for individuals to beat the odds and make it to university.

We further explore how much these non-cognitive skills matter in explaining the socioeconomic gap in university participation using the Kitagawa-Blinder-Oaxaca decomposition. Decomposing the gap in the probability of university participation among the potential FiF and children of graduate parents by gender reveals that the importance of academic self-concept is higher for boys than for girls. Again, this highlights the potential power of interventions that develop academic self-concept, especially for boys from disadvantaged backgrounds. Previous research has made the case that academic self-concept and achievement are mutually reinforcing, leading to gains in a 'virtuous cycle' (Hansen, Henderson, and Shure 2017).

Perhaps surprisingly, from the non-cognitive measures we use in this paper, girls only have an advantage in work ethic while boys on average have higher academic self-concept and higher self-esteem. We believe that the result of boys having higher levels of these two non-cognitive skills than girls is due to the nature of these measures. Academic self-concept and self-esteem are measures of self-beliefs, while work ethic, in which girls exhibit higher levels in this paper, is more related to how people actually behave. Previous research has shown that there are established

differences in how men and women perceive their own abilities (Jerrim, Shure, and Wyness 2020) and as a result decide to behave (Niederle and Vesterlund 2007). This may be especially important in the lead up to high stakes examinations and university applications.

Furthermore, we provide evidence that age 16 test scores are an important channel between adolescent non-cognitive skills and university participation. Once we control for both age 11 and age 16 test scores, age 16 test scores turn out to be more important, suggesting that what happens in school between age 11 and age 16 is critical for university participation. We probe the importance of age 16 test scores by exploring how important our measures of non-cognitive skills are in explaining the socioeconomic gap in these high stakes, compulsory school leaving exams. This analysis shows that academic self-concept is more important than attainment at age 11 for explaining the SES gap in age 16 test scores, especially for boys. Again, this highlights the potential of interventions targeting academic self-concept.

Taken together our results provide evidence that non-cognitive skills, and especially, academic self-concept, are an important facilitator of intergenerational educational mobility. If societies and universities want to improve the diversity of their university cohorts and promote intergenerational educational mobility, then interventions to target the development of these crucial skills should be prioritised. Conducting a meta-analysis, O'Mara et al. (2006) show that targeted interventions can increase academic self-concept by 0.51 standard deviation, which is a meaningful impact.

Our results also underline the importance of early interventions (i.e., well before university application and crucial entrance exams). For policymakers and practitioners interested in developing interventions to target these non-cognitive skills and improve socioeconomic gaps in access to higher education, this may be a critical period to intervene. The findings on gender differences in the importance of non-cognitive skills for intergenerational educational mobility highlight the need to develop interventions that address these differences in a targeted manner. While continuing to improve the academic performance of disadvantaged students should remain a policy priority, focusing on the development of their complementary non-cognitive skills may be just as important in improving intergenerational educational mobility.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge. 2017. 'When Should You Adjust Standard Errors for Clustering?' Working Paper 24003. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w24003.

Adamecz-Völgyi, Anna, Morag Henderson, and Nikki Shure. 2020a. 'Is "First in Family" a Good Indicator for Widening University Participation?' *Economics of Education Review*. https://doi.org/10.1016/j.econedurev.2020.102038.

———. 2020b. 'The Labor Market Returns to "First in Family" University Graduates'. IZA Discussion Papers 13911. Institute of Labor Economics.
Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz. 2011. 'Personality Psychology and Economics'. In *Handbook of the Economics of Education*, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, 4:1–181. Handbook of The Economics of Education. Elsevier. https://doi.org/10.1016/B978-0-444-53444-6.00001-8.
Angrist, Joshua D., and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
Blanden, Jo, Paul Gregg, and Lindsey Macmillan. 2007. 'Accounting for Intergenerational Income Persistence: Noncognitive Skills, Ability and Education'. *The Economic Journal* 117 (519): C43–60. https://doi.org/10.1111/j.1468-0297.2007.02034.x.
Blanden, Jo, and Stephen Machin. 2004. 'Educational Inequality and the Expansion of UK Higher Education'. *Scottish Journal of Political Economy* 51 (2): 230–49.
Blinder, Alan S. 1973. 'Wage Discrimination: Reduced Form and Structural Estimates'. *The Journal of Human Resources* 8 (4): 436–55. https://doi.org/10.2307/144855.
Blundell, Richard, Lorraine Dearden, and Barbara Sianesi. 2005. 'Evaluating the Effect of Education on Earnings: Models, Methods and Results from the National Child Development Survey'. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 168 (3): 473–512. https://doi.org/10.1111/j.1467-985X.2004.00360.x.
Brand, Jennie E., and Yu Xie. 2010. 'Who Benefits Most from College?: Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education'. *American Sociological Review*, April. https://doi.org/10.1177/0003122410363567.
Britton, Jack, Lorraine Dearden, Neil Shephard, and Anna Vignoles. 2016. 'How English Domiciled Graduate Earnings Vary with Gender, Institution Attended, Subject and Socio-Economic Background'. 13 April 2016. https://www.ifs.org.uk/publications/8233.
Card, David. 1999. 'Chapter 30 - The Causal Effect of Education on Earnings'. In *Handbook of Labor Economics*, edited by Orley C. Ashenfelter and David Card, 3:1801–63. Elsevier. https://doi.org/10.1016/S1573-4463(99)03011-4.
Chetty, Raj, John N. Friedman, Emmanuel Saez, and Danny Yagan. 2020. 'Income Segregation and Intergenerational Mobility Across Colleges in the United States'. *The Quarterly Journal of Economics* 135 (3): 1567–1633.
Chowdry, Haroon, Claire Crawford, and Alissa Goodman. 2011. 'The Role of Attitudes and Behaviours in Explaining Socio- Economic Differences in Attainment at Age 16', 18.
Coopersmith, Stanley. 1959. 'A Method for Determining Types of Self-Esteem'. *The Journal of Abnormal and Social Psychology* 59 (1): 87–94. https://doi.org/10.1037/h0048001.
Dearden, Lorraine, Leslie McGranahan, and Barbara Sianesi. 2004. 'Returns to Education for the "marginal Learner" Evidence from the BCS70'. Centre for the Economics of Education. 1 0 2004. https://www.ifs.org.uk/publications/2666.
Department for Education. 2017. 'Statistics: Participation Rates in Higher Education'. Participation Rates in Higher Education for England. 28 September 2017. https://www.gov.uk/government/collections/statistics-on-higher-education-initial-participation-rates.
Dickson, Matt. 2013. 'The Causal Effect of Education on Wages Revisited*'. *Oxford Bulletin of Economics and Statistics* 75 (4): 477–98. https://doi.org/10.1111/j.1468-0084.2012.00708.x.
Fairlie, Robert W. 2005. 'An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models'. *Journal of Economic and Social Measurement* 30 (4): 305–16. https://doi.org/10.3233/JEM-2005-0259.
Gutman, Leslie Morrison, and Ingrid Schoon. 2016. 'A Synthesis of Causal Evidence Linking Non-Cognitive Skills to Later Outcomes for Children and Adolescents'. In *Non-Cognitive Skills and Factors in Educational Attainment*, edited by Myint Swe Khine and Shaljan

Areepattamannil, 171–98. Contemporary Approaches to Research in Learning Innovations. Rotterdam: SensePublishers. https://doi.org/10.1007/978-94-6300-591-3_9.

Hansen, Kirstine, and Morag Henderson. 2019. 'Does Academic Self-Concept Drive Academic Achievement?' *Oxford Review of Education* 45 (5): 657–72. https://doi.org/10.1080/03054985.2019.1594748.

Hansen, Kirstine, Morag Henderson, and Nikki Shure. 2017. 'Does Academic Self-Concept Predict Further and Higher Education Participation?' Centre for Global Higher Education Working Paper 16. Centre for Global Higher Education, UCL Institute of Education. https://www.researchcghe.org/publications/working-paper/does-academic-self-concept-predict-further-and-higher-education-participation/.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. 'The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior'. *Journal of Labor Economics* 24 (3): 411–82. https://doi.org/10.1086/504455.

Henderson, Morag, Nikki Shure, and Anna Adamecz-Völgyi. 2020. 'Moving on up: "First in Family" University Graduates in England'. *Oxford Review of Education*. https://doi.org/10.1080/03054985.2020.1784714.

Jacob, Brian A. 2002. 'Where the Boys Aren't: Non-Cognitive Skills, Returns to School and the Gender Gap in Higher Education'. *Economics of Education Review* 21 (6): 589–98.

Jerrim, John, Nikki Shure, and Gill Wyness. 2020. 'Driven to Succeed? Teenagers' Drive, Ambition and Performance on High-Stakes Examinations'. IZA Discussion Papers 13525. IZA Institute of Labor Economics. https://www.iza.org/publications/dp/13525/driven-to-succeed-teenagers-drive-ambition-and-performance-on-high-stakes-examinations.

Kalil, Ariel, and Salma Khalid. 2010. 'Non-Cognitive Skills and Educational Attainment'. Unpublished manuscript. University of Chicago. https://paa2011.princeton.edu/papers/112079.

Kautz, Tim, James J Heckman, Ron Diris, Bas ter Weel, and Lex Borghans. 2014. 'Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success'. Working Paper 20749. National Bureau of Economic Research. https://doi.org/10.3386/w20749.

Kitagawa, Evelyn M. 1955. 'Components of a Difference Between Two Rates'. *Journal of the American Statistical Association* 50 (272): 1168–94. https://doi.org/10.2307/2281213.

Kraft, Matthew A. 2020. 'Interpreting Effect Sizes of Education Interventions'. *Educational Researcher* 49 (4): 241–53. https://doi.org/10.3102/0013189X20912798.

Mendez, Ildefonso, and Gema Zamarro. 2016. 'Noncognitive Skills'. In *Non-Cognitive Skills and Factors in Educational Attainment*, edited by Myint Swe Khine and Shaljan Areepattamannil, 411–28. Contemporary Approaches to Research in Learning Innovations. Rotterdam: SensePublishers. https://doi.org/10.1007/978-94-6300-591-3_19.

Mendolia, Silvia, and Ian Walker. 2014. 'The Effect of Noncognitive Traits on Health Behaviours in Adolescence'. *Health Economics* 23 (9): 1146–58. https://doi.org/10.1002/hec.3043.

Neumark, David. 1988. 'Employers' Discriminatory Behavior and the Estimation of Wage Discrimination'. *The Journal of Human Resources* 23 (3): 279–95. https://doi.org/10.2307/145830.

Niederle, Muriel, and Lise Vesterlund. 2007. 'Do Women Shy Away From Competition? Do Men Compete Too Much?' *The Quarterly Journal of Economics* 122 (3): 1067–1101. https://doi.org/10.1162/qjec.122.3.1067.

Oaxaca, Ronald. 1973. 'Male-Female Wage Differentials in Urban Labor Markets'. *International Economic Review* 14 (3): 693–709. https://doi.org/10.2307/2525981.

O'Mara, Alison J., Herbert W. Marsh, Rhonda G. Craven, and Raymond L. Debus. 2006. 'Do Self-Concept Interventions Make a Difference? A Synergistic Blend of Construct Validation

and Meta-Analysis'. *Educational Psychologist* 41 (3): 181–206. https://doi.org/10.1207/s15326985ep4103_4.

Oreopoulos, Philip, and Uros Petronijevic. 2013. 'Making College Worth It: A Review of Research on the Returns to Higher Education'. Working Paper 19053. National Bureau of Economic Research. https://doi.org/10.3386/w19053.

Prada, María F., and Sergio Urzúa. 2017. 'One Size Does Not Fit All: Multiple Dimensions of Ability, College Attendance, and Earnings'. *Journal of Labor Economics* 35 (4): 953–91. https://doi.org/10.1086/692477.

Prevoo, Tyas, and Bas ter Weel. 2015. 'The Importance of Early Conscientiousness for Socio-Economic Outcomes: Evidence from the British Cohort Study'. *Oxford Economic Papers* 67 (4): 918–48. https://doi.org/10.1093/oep/gpv022.

Rotter, Julian B. 1966. 'Generalized Expectancies for Internal versus External Control of Reinforcement'. *Psychological Monographs: General and Applied* 80 (1): 1–28. https://doi.org/10.1037/h0092976.

Schurer, Stefanie, Colm P Harmon, Rachael Gibson, and Rebecca Edwards. 2020. 'First in Their Families at University: Can Non-Cognitive Skills Compensate for Social Origin?' IZA Discussion Paper 13721. IZA Institute of Labor Economics. https://www.iza.org/publications/dp/13721/first-in-their-families-at-university-can-non-cognitive-skills-compensate-for-social-origin.

Sloczyński, Tymon. 2020. 'Average Gaps and Oaxaca–Blinder Decompositions: A Cautionary Tale about Regression Estimates of Racial Differences in Labor Market Outcomes'. *ILR Review* 73 (3): 705–29. https://doi.org/10.1177/0019793919874063.

Structural Equation Modeling Reference Manual, Release 16. 2017. Stata Press. https://www.stata.com/bookstore/structural-equation-modeling-reference-manual/.

University College London, UCL Institute of Education, Centre for Longitudinal Studies. 2018. 'Next Steps: Sweeps 1-8, 2004-2016: Secure Access [Data Collection]'. https://doi.org/10.5255/UKDA-SN-7104-4.

Valentine, Jeffrey C., David L. DuBois, and Harris Cooper. 2004. 'The Relation Between Self-Beliefs and Academic Achievement: A Meta-Analytic Review'. *Educational Psychologist* 39 (2): 111–33. https://doi.org/10.1207/s15326985ep3902_3.

Walker, Ian, and Gerda Buchmueller. 2020. 'The Graduate Wage and Earnings Premia and the Role of Non-Cognitive Skills'. *IZA Discussion Paper*, no. No. 13248.

Walker, Ian, and Yu Zhu. 2018. 'University Selectivity and the Relative Returns to Higher Education: Evidence from the UK'. *Labour Economics* 53 (August): 230–49. https://doi.org/10.1016/j.labeco.2018.05.005.

Woessmann, Ludger. 2004. 'How Equal Are Educational Opportunities? Family Background and Student Achievement in Europe and the United States'. 1284. *IZA Discussion Papers*. IZA Discussion Papers. Institute of Labor Economics (IZA). https://ideas.repec.org/p/iza/izadps/dp1284.html.

Appendix A: Questions in Next Steps used to create indexes of non-cognitive skills

1. **External locus of control:** How much do you agree with the following statements? (Potential answers: strongly agree, agree, disagree, strongly disagree)

   Q1: I can pretty much decide what will happen in my life.

   Q2: If you work hard at something, you'll usually succeed.

   Q3: If someone is not a success in life, it is usually their own fault.

   Q4: How well you get on in this world is mostly a matter of luck.

   Q5: People like me don't have much of a chance in life.

   Q6: Even if I did well at school, I would have had a hard time getting the right kind of job.

2. **Academic self-concept**

   Q1: Feelings about school: I get good marks for my work. (Potential answers: strongly agree, agree, disagree, strongly disagree)

   Q2: How good YP thinks YP is at school work? (Potential answers: very good, above average, average, below average, not at all good))

   Q3-Q5: How good or bad at this subject: English, maths, science, and information and communication technology (ICT). (Potential answers: very good, fairly good, not very good, not good at all)

3. **Work ethic** (Potential answers: strongly agree, agree, disagree, strongly disagree)

   Q1: Doing well at school means a lot to me.

   Q2: At school, I work as hard as I can.

   Q3: Working hard at school now will help me to get on later in life.

4. **Self-esteem**

   How useful you have felt recently? (Potential answers: more so than usual, same as usual, less useful than usual, much less useful)

   How much you have been thinking of yourself as a worthless person recently? (Potential answers: not at all, no more than usual, rather more than usual, much more than usual)

Table A1: Data availability on non-cognitive skills by waves in Next Steps

	Wave 1 (age 13)	Wave 2 (age 14)	Wave 3 (age 15)	Wave 4 (age 16)	Missing or no answer
Locus of control					
Q1	6,146				1,390
Q2	6,698				838
Q3	6,213				1,323
Q4	6,260				1,276
Q5	6,392				1,144
Q6	5,953				1,583
Academic self-concept					
Q1	7,115	286	53		82
Q2	7,308				228
Q3	7,434				102
Q4	7,432				104
Q5	7,427				109
Work ethic					
Q1	7,177	271	32		56
Q2		6,667			869
Q3		6,696			840
Self-esteem					
Q1		6,133		6,246	703
Q2		6,641		6,306	607

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016. http://doi.org/10.5255/UKDA-SN-5545-7

Table A2: The correlation matrix of non-cognitive skills

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) External locus of control (PCA)	1						
(2) Academic self-concept (PCA)	-0.307***	1					
(3) Work ethic (PCA)	-0.376***	0.372***	1				
(4) Self-esteem	-0.159***	0.099***	0.074***	1			
Non-cognitive measures via Confirmatory Factor Analysis (CFA)							
(5) External locus of control (CFA)	0.969***	-0.298***	-0.317***	-0.168***	1		
(6) Academic self-concept (CFA)	-0.303***	0.976***	0.380***	0.096***	-0.292***	1	
(7) Work ethic (CFA)	-0.360***	0.321***	0.917***	0.074***	-0.302***	0.326***	1

Notes: N= 4,755. **** p<0.01, ** p<0.05, * p<0.1*

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016. http://doi.org/10.5255/UKDA-SN-5545-7

Appendix B: Confirmatory Factor Analysis with Full Information Maximum Likelihood (FIML)

Table B1: University participation, Confirmatory Factor Analysis estimates (sample of potential FiF)

	PCA (same as Model 5 in Table 3)	CFA model, PCA sample	CFA model, CFA sample
External locus of control	-0.038***	-0.039***	-0.042***
	(0.008)	(0.007)	(0.006)
Work ethic	0.025***	0.029***	0.032***
	(0.008)	(0.008)	(0.006)
Academic self-concept	0.089***	0.087***	0.080***
	(0.009)	(0.009)	(0.007)
Low self-esteem	-0.054**	-0.050**	-0.038**
	(0.023)	(0.023)	(0.018)
Constant	1.208**	1.162**	0.941**
	(0.539)	(0.539)	(0.421)
Observations	3,753	3,753	6,048

Notes: Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Weighted using Wave 8 weights. Control variables: gender, age, number of siblings, region, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, income deprivation, disability, whether their grandparents attended university, Free School Mean (FSM) status, private school, the highest qualification of the mother and father, and Key Stage 2 quintile of achievement.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016: Secure Access. DOI: 10.5255/UKDA-SN-7104-4

Table B2: The relationship between non-cognitive skills and university attendance by KS2 quintiles, Confirmatory Factor Analysis estimates (sample of potential FiF)

	(1) Lowest quintile	(2) Lower-middle quintile	(3) Middle quintile	(4) Upper-middle quintile	(5) Highest quintile
External locus of control	-0.031***	-0.038***	-0.050***	-0.046***	-0.058***
	(0.012)	(0.012)	(0.014)	(0.015)	(0.020)
Work ethic	0.022*	0.019	0.031**	0.052***	0.035*
	(0.013)	(0.016)	(0.014)	(0.015)	(0.020)
Academic self-concept	0.077***	0.082***	0.087***	0.079***	0.055***
	(0.015)	(0.015)	(0.015)	(0.016)	(0.018)
Low self-esteem	-0.001	-0.045	0.052	-0.054	-0.110**
	(0.036)	(0.042)	(0.041)	(0.045)	(0.048)
Constant	1.950**	0.956	0.809	1.051	-0.003
	(0.797)	(0.994)	(0.875)	(1.016)	(1.106)
Observations	1,267	1,173	1,170	1,035	931

Notes: Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Weighted using Wave 8 weights. All models control for gender, age, number of siblings, region, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father,

single parent household, income deprivation, disability, whether their grandparents attended university, Free School Mean (FSM) status, private school, and the highest qualification of the mother and father.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016: Secure Access. DOI: 10.5255/UKDA-SN-7104-4

Table B3: University participation, Confirmatory Factor Analysis estimates, gender*non-cognitive skills interactions (sample of potential FiF)

	PCA (same as Model 5 in Table 3)	CFA model, PCA sample	CFA model, CFA sample
Female	0.087***	0.084***	0.072***
	(0.017)	(0.017)	(0.013)
External locus of control	-0.022*	-0.026**	-0.031***
	(0.012)	(0.010)	(0.009)
Female* External locus of control	-0.037**	-0.029*	-0.023*
	(0.017)	(0.015)	(0.012)
Work ethic	0.022*	0.025**	0.034***
	(0.013)	(0.012)	(0.009)
Female* Work ethic	0.003	0.008	-0.005
	(0.017)	(0.016)	(0.013)
Academic self-concept	0.121***	0.115***	0.098***
	(0.014)	(0.013)	(0.010)
Female*Academic self-concept	-0.062***	-0.056***	-0.035***
	(0.018)	(0.017)	(0.013)
Low self-esteem	-0.034	-0.037	-0.005
	(0.037)	(0.037)	(0.031)
Female* Low self-esteem	-0.026	-0.016	-0.052
	(0.044)	(0.044)	(0.038)
Constant	1.217**	1.191**	0.957**
	(0.540)	(0.542)	(0.421)
Observations	3,753	3,753	6,048

Notes: Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Weighted using Wave 8 weights. Control variables: age, number of siblings, region, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, income deprivation, disability, whether their grandparents attended university, Free School Mean (FSM) status, private school, the highest qualification of the mother and father, and Key Stage 2 quintile of achievement.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016: Secure Access. DOI: 10.5255/UKDA-SN-7104-4

Table B4: University participation by KS2 quintiles, Confirmatory Factor Analysis estimates, gender interactions (sample of potential FiF)

	(1) Lowest quintile	(2) Lower-middle quintile	(3) Middle quintile	(4) Upper-middle quintile	(5) Highest quintile
Female	0.043	0.083***	0.058**	0.133***	0.105***
	(0.028)	(0.029)	(0.029)	(0.030)	(0.034)
External locus of control	-0.020	-0.014	-0.047**	-0.036*	-0.066**
	(0.017)	(0.018)	(0.021)	(0.020)	(0.029)
Female* External locus of control	-0.024	-0.048*	-0.007	-0.013	0.015
	(0.023)	(0.026)	(0.028)	(0.031)	(0.035)
Work ethic	0.020	0.024	0.017	0.049**	0.032
	(0.019)	(0.022)	(0.018)	(0.021)	(0.030)
Female* Work ethic	-0.002	-0.013	0.022	0.003	0.001
	(0.023)	(0.030)	(0.030)	(0.028)	(0.037)
Academic self-concept	0.097***	0.112***	0.123***	0.102***	0.069***
	(0.021)	(0.023)	(0.022)	(0.023)	(0.024)
Female*Academic self-concept	-0.039	-0.053*	-0.067**	-0.031	-0.028
	(0.028)	(0.029)	(0.028)	(0.033)	(0.037)
Low self-esteem	0.007	-0.008	0.041	0.013	-0.045
	(0.058)	(0.063)	(0.064)	(0.079)	(0.088)
Female* Low self-esteem	-0.017	-0.074	-0.003	-0.122	-0.127
	(0.072)	(0.079)	(0.079)	(0.099)	(0.114)
Constant	1.941**	0.497	0.947	0.746	-0.393
	(0.791)	(0.993)	(0.880)	(1.019)	(1.105)
Observations	1,267	1,173	1,170	1,035	931

Notes: Robust standard errors clustered by school in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Weighted using Wave 8 weights. All models control for age, number of siblings, region, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, income deprivation, disability, whether their grandparents attended university, Free School Mean (FSM) status, private school, and the highest qualification of the mother and father.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016: Secure Access. DOI: 10.5255/UKDA-SN-7104-4.

Table B5: Kitagawa-Blinder-Oaxaca decomposition of university attendance among the potential FiF and children of graduate parents: Confirmatory Factor Analysis estimates

	Total sample	Boys	Girls
Group_1: children of graduate parents	0.767***	0.749***	0.783***
	(0.011)	(0.016)	(0.015)
Group_2: potential FiF	0.455***	0.421***	0.481***
	(0.006)	(0.010)	(0.009)
Difference between the two groups	0.312***	0.328***	0.303***
	(0.013)	(0.019)	(0.017)
Explained part	0.202***	0.225***	0.186***
	(0.010)	(0.015)	(0.013)
Unexplained part	0.110***	0.103***	0.117***
	(0.015)	(0.022)	(0.020)
Explained difference			
External locus of control	0.011***	0.011***	0.011***

	(0.002)	(0.003)	(0.002)
Work ethic	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)
Academic self-concept	0.029***	0.040***	0.021***
	(0.003)	(0.005)	(0.004)
Low self-esteem	0.000	0.000	0.000
	(0.000)	(0.000)	(0.001)
KS2	0.019***	0.019***	0.017***
	(0.004)	(0.006)	(0.005)
Family	0.026***	0.025***	0.026***
	(0.004)	(0.006)	(0.005)
Region	-0.021***	-0.031***	-0.001
	(0.005)	(0.007)	(0.002)
Disadvantage	0.037***	0.047***	0.017***
	(0.006)	(0.009)	(0.005)
Birth	-0.017***	-0.013**	-0.015***
	(0.004)	(0.005)	(0.005)
Parents	0.096***	0.108***	0.085***
	(0.009)	(0.013)	(0.012)
Independent school	0.021***	0.019***	0.024***
	(0.004)	(0.006)	(0.005)
Observations	7,507	3,333	4,174

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Weighted using Wave 8 weights. All models control for gender, age, number of siblings, region, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, income deprivation, disability, whether their grandparents attended university, Free School Mean (FSM) status, private school, and Key Stage 2 quintile of achievement. 'Group 1' refers to individuals whose parents have a university degree and 'Group 2' refers to those individuals whose parents do not have a university degree.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016: Secure Access. DOI: 10.5255/UKDA-SN-7104-4

Table B6: Kitagawa-Blinder-Oaxaca decomposition of university attendance with GCSEs and GCSE scores among the potential FiF and children of graduate parents: Confirmatory Factor Analysis estimates

	University participation			GCSE's (standardized)		
	Total	Boys	Girls	Total	Boys	Girls
Group_1: children of graduate parents	0.767***	0.749***	0.783***	0.765***	0.697***	0.826***
	(0.011)	(0.016)	(0.015)	(0.018)	(0.026)	(0.024)
Group_2: potential FiF	0.455***	0.421***	0.481***	0.171***	0.070***	0.250***
	(0.006)	(0.010)	(0.009)	(0.012)	(0.019)	(0.015)
Difference between the two groups	0.312***	0.328***	0.303***	0.593***	0.627***	0.577***
	(0.013)	(0.019)	(0.017)	(0.021)	(0.032)	(0.028)
Explained part	0.250***	0.272***	0.237***	0.413***	0.469***	0.380***
	(0.010)	(0.016)	(0.014)	(0.019)	(0.030)	(0.025)
Unexplained part	0.063***	0.056***	0.066***	0.180***	0.159***	0.196***
	(0.014)	(0.020)	(0.018)	(0.020)	(0.031)	(0.027)
	Explained difference					
External locus of control	0.001	-0.001	0.002	0.041***	0.055***	0.029***
	(0.001)	(0.003)	(0.002)	(0.005)	(0.009)	(0.006)
Work ethic	0.001	0.001	0.001	0.001	0.002	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Academic self-concept	0.010***	0.017***	0.005**	0.078***	0.109***	0.054***
	(0.002)	(0.004)	(0.002)	(0.008)	(0.013)	(0.009)
Low self-esteem	0.000	0.000	0.000	0.001	0.002	0.000
	(0.000)	(0.000)	(0.000)	(0.001)	(0.002)	(0.001)
KS2	0.008**	0.009*	0.006	0.042***	0.044***	0.039***
	(0.003)	(0.005)	(0.005)	(0.006)	(0.009)	(0.008)
GCSE	0.142***	0.144***	0.142***			
	(0.007)	(0.010)	(0.009)			

Family	0.017***	0.018***	0.017***	0.043***	0.041***	0.043***
	(0.004)	(0.006)	(0.005)	(0.006)	(0.009)	(0.008)
Region	-0.013***	-0.019***	-0.002	0.007***	0.012***	0.004
	(0.003)	(0.005)	(0.002)	(0.002)	(0.004)	(0.003)
Disadvantage	0.015***	0.023***	0.002	0.020**	0.019	0.020
	(0.006)	(0.009)	(0.006)	(0.009)	(0.013)	(0.013)
Birth	-0.011***	-0.012**	-0.010**	-0.025***	-0.011	-0.022***
	(0.003)	(0.005)	(0.004)	(0.005)	(0.007)	(0.006)
Parents	0.050***	0.063***	0.040***	0.177***	0.175***	0.176***
	(0.008)	(0.012)	(0.011)	(0.013)	(0.022)	(0.017)
Independent school	0.031***	0.030***	0.032***	0.029***	0.020**	0.036***
	(0.004)	(0.007)	(0.006)	(0.008)	(0.010)	(0.011)
Observations	7,507	3,333	4,174	6,498	2,880	3,618

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All models control for gender, age, number of siblings, region, whether the young person is a care leaver, whether the young person had a low birth weight or was born prematurely, ethnicity, mother's social class, father's social class, age of mother, age of father, single parent household, income deprivation, disability, whether their grandparents attended university, Free School Mean (FSM) status, private school, and Key Stage 2 quintile of achievement. 'Group 1' refers to individuals whose parents have a university degree and 'Group 2' refers to those individuals whose parents do not have a university degree.

Source: University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2018). Next Steps: Sweeps 1-8, 2004-2016: Secure Access. DOI: 10.5255/UKDA-SN-7104-4