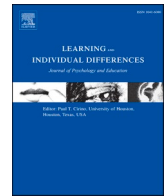




Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Non-cognitive skills and social gaps in digital skills: Evidence from ICILS 2018[☆]

Zbigniew Karpiński^{a,c,*}, Giorgio Di Pietro^{a,d}, Federico Biagi^b

^a European Commission Joint Research Centre, Edificio EXPO, calle Inca Garcilaso 3, 41092 Seville, Spain

^b European Commission Joint Research Centre, Building 26A CCR, via Enrico Fermi 2749, 21027 Ispra, Italy

^c Institute of Philosophy and Sociology Polish Academy of Sciences, ul. Nowy Swiat 72, 00-330 Warsaw, Poland

^d IZA, Schaumburg-Lippe-Strasse 5-9, 53113 Bonn, Germany

ARTICLE INFO

Keywords:

Computer and Information Literacy (CIL)

ICILS 2018

Non-cognitive skills

Gender

Socioeconomic

Native-immigrant gaps in CIL

ABSTRACT

Using data from the 2018 round of the International Computer and Literacy Survey (ICILS), this study looks at the effect of non-cognitive skills (e.g., motivation, ambition, and conscientiousness) on digital competences as measured by the Computer and Information Literacy (CIL) test score. Non-cognitive skills may be especially important in low-stakes tests such as ICILS, where students face no consequences — positive or negative — as a result of their performance. The empirical results show that several non-self-reported measures acting as proxies for non-cognitive skills are significant determinants of CIL test scores. Furthermore, the findings point at differences in non-cognitive skills across gender, immigrant background, and socioeconomic status. This suggests that one should be cautious when inferring about inequality in digital competences along these dimensions using low-stakes test scores, and underscores the importance of controlling for non-cognitive skills.

1. Introduction

Results from large-scale assessments (e.g., PISA, TIMSS, PIRLS) are used for international benchmarking and institutional accountability (Zilberberg et al., 2014). They have also triggered educational policy reforms in several countries (Choi & Jerrim, 2016; Dobbins & Martens, 2012). The rationale behind this is that these test scores — capturing students' knowledge and problem-solving skills in different domains — can be used to compare the efficacy and equity of educational systems around the world. However, several studies (Borghans et al., 2008; Hawthorne et al., 2015) suggest that one needs to be cautious in drawing strong conclusions from such tests as, along with students' cognitive skills, they also reflect their willingness to perform well in the absence of extrinsic reward. Such argument is also supported by the findings of Gneezy et al. (2019) showing that students in various countries come to these assessment tests with different incentives to do well.

Existing research shows that students' performance on a cognitive

test depends not only on the actual cognitive abilities they possess, but also on their non-cognitive skills.¹ While cognitive and non-cognitive skills are correlated, they capture different dimensions with separate effects on students' outcomes (Duckworth et al., 2007; Heckman et al., 2019). Non-cognitive skills have been defined as personality traits that influence the way pupils think, feel and behave (Borghans et al., 2008). Motivation, perseverance, self-control, conscientiousness,² grit³ as well as resilience, creativity, teamwork and social competencies are examples of non-cognitive skills that are highly important for student achievement and predictive of academic success (Gutman & Schoon, 2013; Kattan, 2017; Khine & Areepattamannil, 2016).

Non-cognitive skills may influence student performance particularly on low-stakes tests (i.e., tests that have no impact on students' grades or other academic achievements), such as the international large-scale assessments. While some students may be motivated to perform well irrespective of the rewards offered, the lack of academic consequences for test-takers may discourage others from exerting the maximal effort.

[☆] The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

* Corresponding author at: European Commission Joint Research Centre, Edificio EXPO, calle Inca Garcilaso 3, 41092 Seville, Spain.

E-mail addresses: zbigniew.karpinski@ec.europa.eu (Z. Karpiński), giorgio.di-pietro@ec.europa.eu (G. Di Pietro), federico.biagi@ec.europa.eu (F. Biagi).

¹ Of course, in addition to cognitive and non-cognitive skills, other determinants of standardised test scores include physical, mental, and contextual factors (Rozon, 2013).

² Conscientiousness is a personality trait that reflects the tendency to be responsible, organised, hard-working, and to comply with norms and rules.

³ Grit is a personality trait, which is defined as passion and perseverance to achieve meaningful long-term goals.

<https://doi.org/10.1016/j.lindif.2022.102254>

Received 25 January 2022; Received in revised form 30 November 2022; Accepted 18 December 2022

Available online 31 January 2023

1041-6080/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

We use data from the 2018 wave of the International Computer and Information Literacy Survey (ICILS), a large-scale, international assessment of 8th-grade students' digital competences. We focus on the Computer and Information Literacy (CIL) test, which is expected to reflect "individual's ability to use computers to investigate, create, and communicate in order to participate effectively at home, at school, in the workplace, and in society" (Fraillon et al., 2020b, p. 16). As other large-scale assessments in education, ICILS is a low-stakes test and uses a task-based approach to measuring students' digital skills. Task-based measures of digital skills are preferred over measures based on self-reports, given that the latter are subject to biases and misrepresentations (Juhaňák et al., 2019).

Following recent literature (e.g., Balart et al., 2018; Zamorro et al., 2019), we construct three non-self-reported⁴ measures that act as proxies for students' non-cognitive skills. They are derived from patterns of responses to the CIL test and to the background questionnaire. Specifically, these proxies are: (a) drop in response time to test questions; (b) drop in performance along the test; (c) non-response rate to the background questionnaire. As observed by Borghans and Schils (2018), differences in performance drop during the test are associated with non-cognitive skills such as motivation and ambition. The decline in performance on later portions of the PISA test found in both Borghani and Biecek (2016) and Zamorro et al. (2019) is linked to lower student motivation (see Borghani and Biecek, 2016; Zamorro et al., 2019). Using information from the Programme for the International Assessment of Adult Competencies (PIAAC), Anghel and Balart (2017) argue that time per question is a good indicator for non-cognitive skills such as motivation, ability to stay focused or grit/thrift.⁵ Hedengren and Stratmann (2012) and Hitt et al. (2016) show that non-response to non-challenging questions in surveys is negatively correlated with traits such as conscientiousness and perseverance (see also Zamorro et al., 2020).

This study addresses the following research questions: (1) Do our non-self-reported proxies for students' non-cognitive skills affect student digital competences as measured by the CIL scores? (2) To what extent and in which direction do digital competence gaps along the lines of gender, socioeconomic background, and immigrant status change after accounting for the effect of our non-cognitive skill proxies? The answers to these questions add to the existing literature in two ways. First, our analysis extends the results of earlier studies (e.g., Borghans & Schils, 2018) — which used similar proxies to examine the impact of non-cognitive skills on mathematics, reading, and science test scores — to the area of digital literacy. Indirect evidence suggests that non-cognitive skills may play an important role in predicting ICT-related outcomes. For instance, Grundke et al. (2018) find that digital-intensive industries reward especially workers with high levels of self-organisation (see also Morandini et al., 2020). Second, while data from ICILS 2018 show that females, students from higher socioeconomic status and native students have higher CIL test scores (Fraillon et al., 2020b), we explore the role of unobserved variables such as non-cognitive skills in explaining these

⁴ Non-cognitive skills can be measured using a personality questionnaire designed to classify individuals on dimensions such as conscientiousness or self-control. Alternatively, they can be inferred from patterns of responses (e.g., the number of questions left unanswered) to a questionnaire or test which wasn't designed specifically to measure non-cognitive skills. In the former approach, the relevant indicators are "self-reported", as they are based on information provided directly by students. In the latter approach, the indicators are "non-self-reported", as they are extracted by a researcher. Since the ICILS questionnaire doesn't have any section on personality, we used the non-self-reported approach in this study.

⁵ While, in contrast to ICILS, PIAAC is a test with no time limitation, one can reasonably argue that time per question might be related to non-cognitive skills also in ICILS, where the fixed time that students have to complete the CIL test could affect how they respond to individual questions. For example, more motivated and focused test takers are aware that they cannot spend too much time on a single question if they want to do well on the test.

associations. Several scholars argue that this research is much needed. For instance, in their recent meta-analysis, Scherer and Siddiq (2019) emphasise the importance of exploring additional confounding variables behind the relationship between socioeconomic status and digital skills. The present study addresses precisely this need by looking at the potential contribution of non-cognitive skills.

2. Background and hypotheses

In this section, we first review studies on achievement and non-cognitive skills, focusing on the impact of those non-cognitive skills that the proxies employed in this paper are expected to capture (i.e., motivation, ambition, self-control, ability to stay focused, grit, agreeableness, need for achievement, perseverance, and conscientiousness). In the second part, we look at the literature addressing gaps in digital skills along gender, socioeconomic status, and immigrant background lines and go on to hypothesise how these gaps change when non-cognitive skills are accounted for.

2.1. Non cognitive skills and achievement

Many studies have analysed the importance of non-cognitive traits in student achievement. Lechner et al. (2019) find that non-cognitive skills contribute to the educational outcomes of children and adolescents above and beyond cognitive skills and socioeconomic background. Lee (2020) attempts to identify the most important non-cognitive predictors of student achievement. While Palacios-Abad (2021) concludes that student effort influences the likelihood of completing higher education, Zamorro et al. (2019), using PISA data, demonstrate that this factor explains a considerable part of the variation in test scores across countries.

Moving on to the impact of the specific non-cognitive skills that our proxies attempt to account for, Zigler and Butterfield (1968) provide evidence that motivation significantly improves performance on IQ tests. Similarly, Borghans et al. (2008) find that people with higher motivation generally perform better on cognitive tests. Conscientiousness is also strongly related to academic performance as measured by final grades (Conrad & Patry, 2012) or test scores in mathematics (Meyer et al., 2019).

Tangney et al. (2004) find that students with higher self-control have higher GPAs. Duckworth et al. (2010) conclude that self-control causally impacts academic performance: they find that changes in self-control over time predict variations in GPA.

Using individual-level data from a US university, Duckworth et al. (2007) find that grit scores are associated with higher GPAs and such positive relationship turns out to be robust even when SAT scores are controlled for. This result is in line with that of Strayhorn (2014) who shows that grit influences grades. Hagger and Hamilton (2019) observe that grit is a predictor of science grades in a sample of secondary school students. Thorsen et al. (2021) and Xu et al. (2021) find that perseverance, too, is positively associated with student achievement.

Finally, studies conducted in UK Farsides and Woodfield (2003) and US Bidjerano and Dai (2007) found that agreeableness is positively related to final grades and GPA, respectively.

2.2. Social drivers of digital skills inequalities and the role of non-cognitive skills

2.2.1. The gender dimension

The magnitude and direction of gender differences in digital competencies depends to a large extent on the type of digital skill that is being evaluated. Various studies using data from the first round of ICILS find that female students outperform male students on the CIL test (Gebhardt et al., 2019; Hatlevik et al., 2018; Punter et al., 2017). Upon a closer inspection of the data, Gebhardt et al. (2019) conclude that female students scored higher on tasks involving communication and

creativity, while male students scored higher on tasks requiring some technical skills (see also Guggemos, 2021). Gender gaps in digital skills favouring girls are also found by Aesaert and van Braak (2015) in their study of Flemish grade 6 students. In addition, meta-analysis results by Siddiq and Scherer (2019) indicate a positive and statistically significant gap in digital skills in favour of female students. Using a large longitudinal sample of German students, Gnamb (2021) tracks changes in their digital skills and digital confidence over the course of three years, i.e., between the ages of 15 and 18. He shows that, while gender gaps favouring male students were initially negligible, they grew considerably over the period under study.

As for non-cognitive skills, several studies indicate that, on average, women show higher motivation, especially in low-stakes tests. DeMars et al. (2013) show consistent evidence supporting the proposition that men exhibit lower test-taking motivation than women (Di Chiacchio et al., 2016; Yunus & Ali, 2009, see also). Additionally, female students score higher on both conscientiousness (Schmitt et al., 2008) and agreeableness (Costa et al., 2001; Feingold, 1994). Wehner and Schils (2021) argue that conscientiousness levels tend to be higher among females relative to males, particularly with respect to aspects such as better organisation and diligence. Based on data from a nationally representative sample of American students, DiPrete and Jennings (2012) find that teachers consistently rate girls higher than boys on various non-cognitive skills, including self-control. Using data from PISA 2012, Arellano et al. (2018) look at the effect of non-cognitive skills on test scores. They find that, while females have higher scores than males, accounting for perseverance and motivation reduces the gender gap.

In light of the results of previous studies, the following hypothesis is made:

Hypothesis 1. *Assuming that girls outperform boys on the CIL test, accounting for differences in non-cognitive skills between the former and the latter reduces the size of the gender gap in the CIL scores.*

2.2.2. The socioeconomic dimension

Existing studies clearly indicate that individuals from more advantaged backgrounds are more likely to be digitally skilled, on average, than those from less advantaged backgrounds. This finding seems to hold regardless of the specific variable representing socioeconomic status (mother's education, see Aesaert & van Braak, 2015; a composite index based on parental education, parental occupation, and household possessions, see Hatlevik et al., 2018; cultural capital, see Hatlevik, Guðmundsdóttir, et al., 2015; Hatlevik, Ottestad, et al., 2015).

There is also a large consensus that socioeconomic status is an important determinant of students' non-cognitive skills (see, for example, Liu, 2019) and especially motivation. Children from higher status families are more likely to develop an academic intrinsic motivation than those from lower status families (Farkas, 2003). Two reasons are often advanced in an attempt to explain such finding. First, parents' beliefs and expectations may exert a significant influence on their children's intrinsic motivation (Bandura, 1997). Less educated parents are less likely to place a high value on their children's education and hence are less inclined to transmit them a positive attitude towards learning and their academic abilities (Gottfried et al., 1994; Wang et al., 2020). As a result of this, students from more advantaged backgrounds are more likely to exhibit an autonomous learning behaviour than their peers from less advantaged backgrounds (Kormos & Kiddle, 2013). Kormos and Csizér (2014) indicate that autonomous learning and effective self-regulated strategies are related to student motivation. Second, socioeconomic status may influence the type of activities children do outside home. For instance, children of more educated parents are more likely to participate in extracurricular activities that stimulate academic mindset and motivation (Usher & Kober, 2012). Using 2015 PISA data, Azzolini et al. (2019) examine the relationship between student perseverance during the test and socioeconomic status. They find that students from advantaged backgrounds have more perseverance than their less

advantaged peers.

Based on these findings, the following hypothesis is proposed:

Hypothesis 2. *Assuming that students from more advantaged backgrounds outperform students from less advantaged backgrounds on the CIL test, accounting for differences in non-cognitive skills between the former and the latter reduces the size of the socioeconomic gap in the CIL scores.*

2.2.3. The immigrant background dimension

Immigrant students have, on average, lower academic achievements, on than their native peers. This holds in general and also for digital skills (Fraillon et al., 2020b, p. 83). In spite of this, students with an immigrant background tend to report higher levels of motivation (Alivernini et al., 2018; Areepattamannil & Freeman, 2008; Schleicher, 2006). Two explanations can be proposed. First, the higher academic motivation of immigrant students may stem from a sense of family obligation (Fuligni & Tseng, 1999). They may feel obliged to do well at school in order not to disappoint their parents (Urdan, 2004), many of whom have made enormous sacrifices by leaving their country of origin in search for a better life for themselves and their children. Educational success is perceived by these students as a way to repay their parents for their sacrifices and investment and financially assist their families in the future (Tseng, 2002). Second, it is also possible that students from immigrant backgrounds are autonomously motivated to learn (Alivernini et al., 2018). They are aware that education is an important asset that will help integrate them successfully into the host country both economically and socially (St-Hilaire, 2002). Individuals with a higher level of education have better job opportunities and higher incomes. Education also plays a crucial role in helping immigrants assimilate the culture of the host country.

Using data on 11th-grade and 12th-grade students from two public secondary schools in Great Toronto Area, Areepattamannil and Freeman (2008) find that immigrant adolescents report higher intrinsic and extrinsic motivation in comparison with their non-immigrant counterparts. Schleicher (2006) uses PISA 2003 data to look at differences in learning dispositions between immigrant students and their native peers. He finds that the former report higher levels of motivation in mathematics than the latter. Employing data from a representative sample of Italian 5th-grade students, not only do Alivernini et al. (2018) find that immigrant students are more academically motivated than natives, but they also conclude that first generation immigrant pupils are more motivated and engaged in the school context than their second generation counterparts.

Our final hypothesis is:

Hypothesis 3. *Assuming that foreign-born students fare worse on the CIL test than their native-born peers, accounting for differences in non-cognitive skills between the former and the latter widens the immigrant-native gap in the CIL scores.*

3. Data and variables

This study uses data from the second round of the ICILS survey, conducted in 2018. While most ICILS participants are countries, a city (i.e., Moscow) and a region (i.e., North Rhine-Westphalia) have also taken part in this survey. The sample is designed to be representative of the relevant population of 8th-grade students. While, in addition to students, ICILS collects also information on teachers and headmasters related to the schools attended by the students, these data are not used in our analysis.

Overall, sample sizes of student participants in the second round of ICILS range from 1991 in North Rhine-Westphalia to 6790 in the United States, and median sample size is equal to 2908. However, due to missing values on key variables, which we listwise-deleted, the analysis is based on somewhat reduced samples. The number of missing cases varies by country/region/city, ranging from 112 in Denmark to 1394 in Luxembourg. The dataset that we finally use for all the analyses reported

in this paper comprises 40403 students. All individuals whose data were collected in the survey used in this study provided consent for publication of these data (Fraillon et al., 2020a, chap. 8).

We use the following student characteristics as independent variables in our analysis: gender (coded 1 for female, and 0 for male), age (in years), immigrant background (captured by a binary variable that takes the value of 1 if both parents were born abroad, and 0 if at least one parent was born in the country of the test), the number of computers (i. e., desktop computers or laptops) at home (captured by a binary variable that takes the value of 1 if there are at least two computers, and 0 otherwise), experience using computers (dummy variables are created for the following categories: never or less than one year; at least 1 year but less than 3 years; at least 3 years but less than 5 years; at least 5 years but less than 7 years; and 7 years or more), parental education (captured by a binary variable coded 1 for students who have at least one parent with tertiary education, and 0 otherwise). Note that the information on all the student characteristics, including parental education, is self-reported by students.

3.1. The design of the ICILS test

An important feature of the ICILS test lies in its “rotated design”: different versions of the test are developed and then allocated randomly among participants. The versions differ in terms of content (i.e., some tasks are included in some versions, but not in others) and ordering of the tasks (i.e., the same task can be placed early or late in the test question sequence, depending on the version). This latter characteristic allows for studying the effect of item position, which is exogenously varied, on students’ behaviour (e.g., response time or the probability of answering correctly).

Digital skills are a latent ability that is not directly observable and is estimated with considerable uncertainty. To acknowledge the uncertainty, students are assigned a set of five “plausible values” to represent their performance on the test (Khorramdel et al., 2020; Lechner et al., 2021). These plausible values are scaled to have a mean of 500 and a standard deviation of 100. Analyses are performed on each plausible value separately and the results of these analyses are then aggregated following Rubin’s rules (Drechsler, 2020) to generate point estimates of the parameters of interest and their standard errors. Table 1 shows the average scores on the CIL test, along with the corresponding standard errors, for each country/region/city participating in ICILS 2018. Students’ performance on the test, as represented by the plausible values, is our dependent variable in the analysis.

Before presenting the results of the analysis, it is important to understand how the ICILS test was organised. Overall, five different CIL modules were developed, each lasting up to 30 min. An individual test comprised two of these modules, so that a total of 20 different versions of the CIL test could be obtained. These versions were allocated

Table 1

Average CIL test scores, standard deviations and sample sizes, by participating country/region/city.

Country/region/city	Mean	S.D.	Sample size
Denmark	552.6	65.5	2404
Moscow	549.1	68.0	2852
Korea	542.1	94.9	2875
Finland	530.7	80.7	2546
United States	518.9	80.7	6790
Germany	518.3	80.1	3655
Portugal	516.5	71.1	3221
North Rhine-Westphalia	514.7	75.1	1991
France	498.7	80.5	2940
Luxembourg	481.8	83.7	5401
Chile	475.9	83.3	3092
Italy	461.0	81.7	2810
Uruguay	450.4	100.5	2613
Kazakhstan	395.2	106.0	3371

randomly to students. They had an hour to complete the test, after which they filled in the background questionnaire.

3.2. Proxies for non-cognitive skills

Along with the student characteristics described above, this study uses three proxies for non-cognitive skills. The first proxy is the drop in response time to successive questions in the test. As mentioned in the introduction, the average time per question is indicative of such non-cognitive skills as the ability to stay focused or grit (Anghel & Balart, 2017). By implication, a decline in the response time over the course of the test can be interpreted in terms of lack of these skills. Given the rotated test design, we can measure a decline in the response time during the test net of the characteristics of the questions (e.g., difficulty or complexity).

To measure drop in response time, we adapt the method developed by Borghans and Schils (2018). We assign a sequence number to each question in a given version of the test and go on to rescale the numbers so that the first question has number 0 and the last question has number 1. Then we regress response time on the rescaled question/sequence number. The coefficient for the latter variable represents the average difference in response time between the last and the first question on the test, regardless of the test version. We use mixed-effect linear regression with crossed random effects associated with respondents and with questions to account for a “grouped” character of our data: each student answers more than one question and the same question is answered by more than one student, so the observations can be said to be “nested within” respondents and test questions. By fitting a mixed-effect model with varying slopes, we can extract student-specific slopes from the model and use them as a variable representing drop in response time.

Our second proxy of non-cognitive skills is the drop in the probability of a correct answer to each successive question on the test. Following the literature reviewed above, this variable is associated with such non-cognitive skills as self-control (Borgonovi & Biecek, 2016), motivation and ambition (Borghans & Schils, 2018), conscientiousness, agreeableness and the need for achievement (Borghans & Schils, 2018). Our approach to measuring performance drop consists in modelling the probability of a correct answer to a given question as a function of its placement in the test.

Building on the original work by Borghans and Schils (2018), we retrieve performance drop from students’ responses to test questions while following the same steps described above for drop in response times. We regress responses to the test questions on the (rescaled) question/sequence number using mixed-effect linear model with crossed random effects associated with respondents and with questions. Correct responses are coded as 1 and incorrect responses as 0. Responses with partial credit are coded as 0.5. We fit a random-intercept and random-slope model, where the slope is allowed to vary across students, and then we extract student-specific slopes to use them as proxies for the above-mentioned non-cognitive skills. The slope coefficient for the (rescaled) question number is the difference in the probability of a correct response between the last and the first question in the test. Again, because of the rotated test design used in ICILS, the probability is estimated net of characteristics of the questions which also affect the probability of a correct response. Negative values of the variable indicate lower levels of the non-cognitive skills that this variable represents, while its positive values suggest higher levels of those skills.

Our third proxy for non-cognitive skills derives from counting the number of times students provide meaningful responses to questions in the background questionnaire. In some cases, students leave out a question without answering. Non-response, or a failure to answer questions in the background questionnaire can indicate a low level of conscientiousness (Hedengren & Stratmann, 2012; Hitt et al., 2016; Soland et al., 2019). Subtracting the number of unanswered questions from the total number of questions in the background questionnaire gives us the number of meaningful responses provided by a student. By

implication, this number is expected to be positively associated with conscientiousness. For the ease of interpretation, we divide it by the total number of questions, so that the resulting values are bounded between 0 and 1, with higher values indicating higher levels of conscientiousness. In the text, we refer to this variable as “item response rate”.

In [Table 2](#), we show pairwise correlations between the variables used in the current study. In most cases, correlations are rather moderate. Specifically, the correlations between our three proxies for non-cognitive skills are not very strong. This confirms that, rather than reflecting the same underlying construct, the three variables represent distinct and separate constructs.

4. Methods

We use linear models in which CIL plausible values are regressed against the independent variables identified in Section 3. Recall that the hypotheses we wish to test are that our proxies for non-cognitive skills act towards narrowing the CIL gaps along gender and socioeconomic status lines, while they work in the opposite direction as regards the immigrant background gap. Testing these hypotheses requires a comparison of a “baseline” model, which doesn’t include proxies for non-cognitive skills, with an “extended” model where our proxies for non-cognitive skills are added. If such hypotheses are correct, then the regression coefficients for gender and socioeconomic status will be statistically significantly smaller in the extended model than in the baseline model, whereas the opposite will occur for the coefficient on immigrant background.⁶

It is reasonable to assume that some part of the variability in CIL scores can be attributed to systematic differences between schools in terms of ICT-related resources and policies. That is, some schools are likely to give their students more opportunities than others to develop their digital skills. It is of paramount importance to account for the influence of these factors as they are likely to be related to some of the variables that are of interest in this study. For example, if students from more advantaged backgrounds are more likely to end up in schools better endowed with IT resources than students from less advantaged backgrounds, then the coefficient on socioeconomic status will be picking up the effect of the between-school differences in ICT resources. Thus, unless accounted for, the between-school variability may result in an upwardly biased estimate of the coefficient on socioeconomic status.

To handle this problem, we mean-center all our variables at the school-level. This way, we are able to remove the part of the variability in CIL test scores that is due to schools — and also, indirectly, the one due to countries/regions/cities — leaving only the part that can be attributed to differences across individual students ([Enders & Tofghi, 2007](#)). In other words, both the between-school and between-country/region/city variances of the mean-centred variables are zero.⁷

Since the dataset used in this paper is organised along three levels, with students nested in schools and schools nested in countries/regions/cities, a multilevel framework could, in principle, be used in the analysis of ICILS 2018. However, given that the data come from a complex, multistage survey, they have to be properly weighted at each level of the

⁶ Testing our hypotheses amounts to testing if the difference between regression coefficients on a given variable from two different regression models is statistically significantly different from 0. Performing such a test requires that the standard error of the difference is calculated first. Calculating the standard error, however, is not trivial ([Clogg et al., 1995](#); [Mize et al., 2019](#)). In this study, we take advantage of replicate weights to estimate the standard error of a difference between two regression coefficients.

⁷ A school-specific mean of a variable mean-centred the school level is equal to 0. This applies to all schools in a given country/region/city, so the average taken over all the schools within that country/region/city will also be equal to 0. Thus, since the mean-centred variable has a mean of 0 in each country/region/city, there is no between-country/region/city variance.

analysis to obtain unbiased estimates of the relevant parameters. Unfortunately, while survey weights for students and schools are provided in the ICILS 2018 dataset, this does not occur for countries/regions/cities. This ultimately reflects the fact that countries/regions/cities were not randomly selected from an underlying population. One therefore faces a choice between (a) fitting a three-level model to the pooled data discarding the survey weights altogether and (b) fitting a two-level model with survey weights for students and schools for each of the participating countries/regions/cities separately. In the former case, however, the resulting parameter estimates are likely to be biased. In the latter case, in turn, an additional difficulty arises: a rather high rate of missing data at the school level. The fixed-effects approach, by contrast, allows one to run the analysis on the pooled dataset using only student-level data (including the survey weights). This way, one can obtain unbiased estimates of the parameters of interest while controlling for time-invariant characteristics, observed and unobserved, of schools and countries/regions/cities (see [Jäger, 2011](#), for an example of a similar approach).

All the regressions reported in the next section are estimated using the mean-centred variables. As a result, the intercept in all the models is equal to 0. However, the slope coefficients have the same interpretation they would have in a standard OLS model. That is, they can be interpreted in terms of a change in the CIL test score associated with a 1-unit change in the given independent variable when all other independent variables are held constant. Because ICILS is based on a complex sampling design, point estimates of the model parameters are obtained by applying survey design weights. Standard errors of the estimates, in turn, are obtained using replicate weights included in the ICILS dataset (for details concerning the sample design and how to use the weights in the analyses of ICILS data, see [Mikheeva & Meyer, 2020](#)). All the regression models are estimated using functions in the R package `survey` ([Lumley, 2010, 2020](#)). Note that the standard errors returned by the package are heteroskedasticity-consistent by default.

5. Results

To verify whether our proxies for non-cognitive skills vary systematically by gender, socioeconomic status, and immigrant background, we regressed these proxies against the set of independent variables described earlier. The results, which are presented in [Table 3](#), show that the coefficient on gender is consistently positive and statistically significant, confirming that female students outperform male students in each of the non-cognitive skills. As for socioeconomic status, its coefficient is positive across all three models, as expected. However, it fails to reach statistical significance in the model for item response rate. Finally, against our expectations, the coefficient on immigrant background is negative, suggesting that migrant students tend to have lower levels of the relevant non-cognitive skills than their native-born peers. The fact that the difference in non-cognitive skills by immigrant background is in the opposite direction of what we expected has important implications for our [Hypothesis 3](#). We will return to this issue in the next Section.

Along with the regression coefficients, [Table 3](#) provides (in italics) values of partial (Cohen’s) f^2 , a common measure of effect size ([Selya et al., 2012](#)). The f^2 values of 0.35, 0.15, and 0.02 are used as thresholds for large, medium, and small effect sizes, respectively. Compared to these thresholds, the regression coefficients in [Table 3](#), even if statistically significant, have very small effect sizes. Overall, the independent variables account for only a small part of the variability in our proxies for non-cognitive skills, which is further confirmed by low values of R^2 and the adjusted R^2 (note that, unlike R^2 which is always bounded between 0 and 1, adjusted R^2 can assume negative values).

Moving on to the main focus of our analysis, [Table 4](#) presents results for another set of regression models. Model 1 is our baseline model. In Models 2 through 4 we add our proxies for non-cognitive skills, one at a time. In Model 5, all three proxies are included together. Using VIF to check for multicollinearity, we find that all variables included in our

Table 2

Pairwise correlations between the variables used in the study; ICILS 2018, pooled sample. NB: For computer experience, the reported values refer to Kendall's τ rank correlation coefficient.

	Female	Age	Migrant	Computer experience	Computers at home	Parent with higher education	Time drop	Performance drop	Response rate	CIL test score
Female		-0.06***	0	0.01*	0.02**	-0.03***	0.13***	0.11***	0.05***	0.12***
Age			0.04***	0.06***	-0.03***	0	-0.09***	-0.02**	-0.03***	-0.02
Migrant				-0.02*	-0.04***	-0.04***	0	-0.04***	-0.05***	-0.09***
Computer experience					0.2***	0.15***	0.05***	0.06***	0.06***	0.18***
Computers at home						0.2***	0.09***	0.09***	0.06***	0.24***
Parent with higher education							0.06***	0.09***	0.06***	0.23***
Time drop								0.34***	0.14***	0.53***
Performance drop									0.12***	0.39***
Response rate										0.28***
CIL test score										

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table 3

Results from for fixed-effect linear models for the three proxies for non-cognitive skills.

	Time drop	Performance drop	Item response rate
Female	11.34*** (0.80) <i>0.018</i>	0.02*** (0.00) <i>0.012</i>	0.01*** (0.00) <i>0.003</i>
Age	-5.21*** (0.88) <i>0.003</i>	-0.01*** (0.00) <i>0.002</i>	-0.01** (0.00) <i>0.001</i>
Migrant	-3.96** (1.39) <i>0.001</i>	-0.01** (0.00) <i>0.001</i>	-0.01* (0.00) <i>0.000</i>
Computers at home	1.29 (0.89) <i>0.000</i>	0.01** (0.00) <i>0.000</i>	0.01** (0.00) <i>0.001</i>
Computer experience: [1, 3]	5.83*** (1.27) <i>0.001</i>	0.01* (0.00) <i>0.000</i>	0.01 (0.00) <i>0.000</i>
Computer experience: [3, 5]	6.73*** (1.55) <i>0.002</i>	0.01*** (0.00) <i>0.001</i>	0.01* (0.00) <i>0.001</i>
Computer experience: [5, 7]	9.36*** (1.33) <i>0.003</i>	0.02*** (0.00) <i>0.001</i>	0.01* (0.00) <i>0.001</i>
Computer experience: 7+	9.30*** (1.29) <i>0.003</i>	0.01*** (0.00) <i>0.001</i>	0.02*** (0.00) <i>0.002</i>
Parents with higher education	3.72*** (0.90) <i>0.001</i>	0.01** (0.00) <i>0.001</i>	0.001 (0.00) <i>0.000</i>
N	40,403	40,403	40,403
R ²	0.03	0.02	0.01
Adj. R ²	-0.11	-0.12	-0.13

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Values in italic font indicate effect sizes for the regression coefficients.

models have VIF values well below 10, which means that there is no strong multicollinearity among the variables considered (see Gujarati, 2004, p. 362).

The inclusion of the proxies for non-cognitive skills considerably improves the explanatory power of the model, as indicated by the values of the adjusted R² in Table 4. While drop in response time turns out to have the largest effect on the adjusted R², item response rate is found to have the lowest impact. This could be attributed to the relatively small variance of this variable in the ICILS 2018 sample.⁸

The signs of the coefficients on the proxies for non-cognitive skills

⁸ The distribution of this variable is highly skewed, with most students having values close to 1.

are consistent with our expectations. Since drop in response time refers to a change in response time between the last and the first task, a value of -60 means that a student takes 60 s less to answer the last question in the test in comparison with the first one. On the other hand, a value of 30 indicates that a student takes 30 s more to solve the last task than the first one. Thus, following the interpretation of response time offered in previous studies, the former student is seen as less able to stay focused or lacking in grit than the latter. According to the estimates in Table 4, the difference in drop in response time between the two aforementioned students predicts a difference of $[30 - (-60)] \cdot 0.78 = 70.2$ CIL points in Model 2 and of $[30 - (-60)] \cdot 0.67 = 60.3$ CIL points in Model 5, holding the other independent variables constant.

Similarly, performance drop corresponds to a change in the probability of a correct answer between the last and the first question in the test. Thus, a value of -0.3 means that the probability of a correct answer drops by 30 % in the course of the test. By the same token, a value of 0.2 means that the probability goes up by 20 %. A student with the latter score is perceived to be more motivated and ambitious, more agreeable and conscientious and having a greater need for achievement than the former. The aforementioned difference in performance drop predicts a gap of $[0.2 - (-0.3)] \cdot 208.56 = 104.28$ CIL points in Model 3 and of $[0.2 - (-0.3)] \cdot 117.65 = 58.83$ CIL points in Model 5, keeping the other independent variables constant.

As for the item response rate, its coefficient can be interpreted as the expected difference in test scores between a student who answered all the questions in the background questionnaire and a student who answered none. For example, a student with a value of 0.9 answered 90 % of all the questions in the background questionnaire. Similarly, a person with a value of 0.75 answered three quarters of the questions. According the estimates reported in Table 4, this difference results in a gap of $(0.9 - 0.75) \cdot 157.7 = 23.66$ CIL points in Model 4 and of $(0.9 - 0.75) \cdot 119.17 = 17.88$ CIL points in Model 5, keeping the other independent variables constant.

As before, partial (Cohen's) f^2 for each estimated coefficient is reported (in italics) in Table 4. For the three variables of interest — gender, socioeconomic status, and immigrant background, the effect sizes are pretty small, whereas for the proxies for non-cognitive skills they turn out to be small to medium, depending on the model specification, with the largest effect sizes observed for time drop. The small effect sizes should be a reason for caution when interpreting the regression coefficients, since, in spite of being statistically significant, they may not be “practically significant”. Such a result is not uncommon in studies using data from large-scale assessments in education, as shown in the online supplement to this paper.

Fig. 1 presents the results of the tests of our hypotheses. For each characteristic of interest, we compare its coefficient in the baseline model (Model 1) with the corresponding coefficient in the extended models (Models 2 through 5). Labels on the horizontal axis indicate the

Table 4
Results from the fixed-effects linear models for the CIL test scores.

	Model 1	Model 2	Model 3	Model 4	Model 5
Female	16.91*** (1.37) <i>0.016</i>	8.01*** (1.30) <i>0.005</i>	12.30*** (1.36) <i>0.009</i>	15.19*** (1.31) <i>0.013</i>	5.39*** (1.27) <i>0.002</i>
Age	-16.30*** (1.46) <i>0.011</i>	-12.21*** (1.30) <i>0.008</i>	-14.39*** (1.41) <i>0.010</i>	-15.16*** (1.39) <i>0.011</i>	-10.86*** (1.22) <i>0.007</i>
Migrant	-14.05*** (2.41) <i>0.003</i>	-10.94*** (2.06) <i>0.002</i>	-12.11*** (2.24) <i>0.002</i>	-12.74*** (2.35) <i>0.002</i>	-9.31*** (1.97) <i>0.002</i>
Computers at home	11.97*** (1.31) <i>0.006</i>	10.96*** (1.18) <i>0.006</i>	10.90*** (1.19) <i>0.005</i>	11.07*** (1.27) <i>0.005</i>	9.82*** (1.12) <i>0.006</i>
Computer experience: [1, 3)	21.26*** (2.53) <i>0.006</i>	16.68*** (2.35) <i>0.005</i>	19.70*** (2.40) <i>0.006</i>	20.44*** (2.51) <i>0.006</i>	15.84*** (2.27) <i>0.005</i>
Computer experience: [3, 5)	33.03*** (2.83) <i>0.015</i>	27.75*** (2.61) <i>0.014</i>	30.41*** (2.63) <i>0.014</i>	31.39*** (2.79) <i>0.014</i>	25.78*** (2.54) <i>0.013</i>
Computer experience: [5, 7)	41.06*** (2.69) <i>0.022</i>	33.72*** (2.40) <i>0.019</i>	37.78*** (2.47) <i>0.020</i>	39.38*** (2.74) <i>0.021</i>	31.66*** (2.31) <i>0.018</i>
Computer experience: 7+	44.68*** (2.51) <i>0.028</i>	37.38*** (2.25) <i>0.025</i>	41.69*** (2.33) <i>0.027</i>	42.07*** (2.40) <i>0.026</i>	34.77*** (2.09) <i>0.024</i>
Parents with higher education	12.12*** (1.36) <i>0.006</i>	9.20*** (1.17) <i>0.005</i>	10.98*** (1.27) <i>0.006</i>	11.93*** (1.30) <i>0.006</i>	8.83*** (1.12) <i>0.005</i>
Time drop		0.78*** (0.02) <i>0.321</i>			0.67*** (0.02) <i>0.234</i>
Performance drop			208.56*** (5.29) <i>0.109</i>		117.65*** (5.23) <i>0.041</i>
Item response rate				157.70*** (6.61) <i>0.060</i>	119.17*** (6.14) <i>0.046</i>
<i>N</i>	40,403	40,403	40,403	40,403	40,403
<i>R</i> ²	0.09	0.31	0.18	0.14	0.37
Adj. <i>R</i> ²	-0.04	0.20	0.05	0.00	0.24

****p* < 0.001.; ***p* < 0.01.; **p* < 0.05

Values in italic font indicate effect sizes for the regression coefficients.

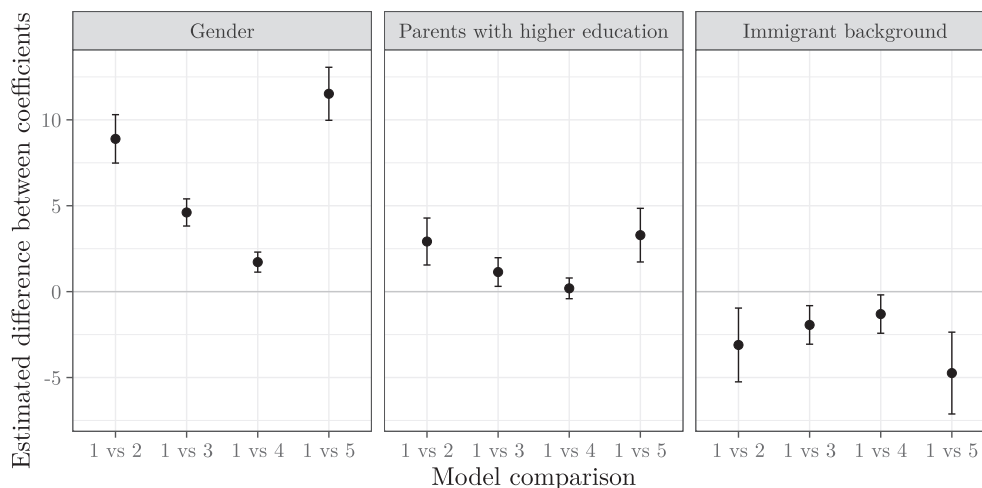


Fig. 1. Differences in estimates of the coefficients on female, parents with tertiary education, and immigrant background and their corresponding 95 % confidence intervals across the models shown in Table 4.

models that are being compared. The grey dots in Fig. 1 represent point estimates of the difference between the relevant regression coefficients in the models under comparison, whereas the vertical segments represent 95 % confidence intervals around the difference. Finally, the horizontal line at 0 represents the null hypothesis of no difference in the

value of the coefficients in the two models under comparison. The coefficient on gender in the baseline model equals 16.91 (i.e., female students outperform male students by about 17 % of the standard deviation). However, adding drop in response time reduces the gender score gap to 8.01, a difference of nearly 9 points that is statistically

significantly different from 0 at the level of $p = 0.05$ (see left panel of Fig. 1, the comparison of Models 1 and 2). When performance drop is added to the model, the estimated regression coefficient on gender is equal to 12.30, a difference of 4.61 points relative to the baseline model. Again, this difference is statistically significant. A similar observation can be made regarding the comparison of the gender coefficients in Models 1 and 4, although the difference in this case is rather small and, while statistically significant, has little practical meaning. Finally, the gender coefficient in Model 5 (where all the three proxies for non-cognitive skills are simultaneously included) is found to be equal to 5.39, a difference of more than 10 points relative to the baseline model, implying that a considerable part of the gender gap in the CIL test scores can be attributed to differences between male and female students in terms of non-cognitive skills. All in all, all our proxies turn out to narrow the gender gap in CIL scores, thus lending support to [Hypothesis 1](#).

A similar pattern is observed as regards the effect of our proxies for non-cognitive skills on the socioeconomic gap in digital competences. The results of our baseline model suggest that students who have at least one parent with higher education score 12.12 CIL points more than those whose parents don't have higher education. In Model 2 which, along with all the variables included in the baseline model, also includes time drop, the estimated coefficient on parental education is lower by 3 points, a statistically significant difference (see Fig. 1, the comparison of Models 1 and 2 in the middle panel). To sum up, adding the proxies for non-cognitive skills to the baseline model leads to a statistically significant reduction in the socioeconomic gap (see Fig. 1, the middle panel), thereby lending support to [Hypothesis 2](#), in all models except for Model 4.⁹

Finally, students with an immigrant background are, on average, found to have a lower CIL test score than their otherwise comparable native peers. While in the baseline model the estimated difference between native and immigrant students is about 14 points, in the full model the corresponding figure is 9.31 points. Thus, including all the three proxies reduces the gap by a third. What is more, the difference in the estimated coefficient on migrant background is statistically significant for all model comparisons, which is *not* consistent with [Hypothesis 3](#).

Coefficients on computer experience and the number of computers at home are in line with expectations. As for age, its coefficient is negative and statistically significant, implying that, all else being equal, older students score worse on the CIL test. This mirrors results from other studies, not limited to student populations (see [Munger et al., 2021](#), for instance), in which age was also found to correlate negatively with digital skills.

6. Discussion and conclusion

Following an increasing emphasis worldwide on the need to constantly monitor and compare educational quality and student learning across countries, there has been an expansion of international large-scale assessments. A great deal of evidence (e.g., [Breakspear, 2012](#); [Lundahl & Serder, 2020](#)) suggests that the outcomes of these tests are taken in great consideration by educational stakeholders including policymakers, parents, and teachers. However, some students may not perform to the best of their ability given that such tests have no impact on their grades or other academic pursuits. Importantly, there might be systematic differences in this respect not only across students within a country, but also across students from different countries. All this underscores the importance of accounting for non-cognitive skills (e.g., motivation, ambition, and conscientiousness) before any inference is

⁹ Note that all of the above conclusions hold when, instead of parental education, we use other proxies for students' socioeconomic status (the results are available upon request). This suggests that our findings are robust with respect to the variable used to capture socioeconomic status.

made from the results of these tests, especially in an international comparison setting.

Building on the findings of previous studies showing that non-cognitive skills affect mathematics, science, and reading test scores ([Borghans & Schils, 2018](#)), we conclude that the same holds also for digital skills. Using data from ICILS 2018, we find that different non-self-reported measures acting as proxies for non-cognitive skills turn out to be significant predictors of student digital competences as measured by the CIL test scores. More specifically, drop in response time to test questions, drop in performance along the test and non-response rate in the background questionnaire are all related to students' CIL test scores. Drop in response time, an indicator for student motivation, ambition, conscientiousness and agreeableness, emerges as the strongest predictor of student achievement among our proxies for non-cognitive skills.

Additionally, the empirical findings point at differences in non-cognitive traits across gender, immigrant background and socioeconomic status. Females, natives and students from more advantaged backgrounds tend, on average, to exhibit higher non-cognitive skills (as captured by our three proxies) than males, immigrants and students from less advantaged backgrounds, respectively. Our result on gender is consistent with expectations because, as argued earlier, women tend to be more conscientious and more motivated than men in schoolwork. In line with this consideration, it is not surprising that previous studies (see, for instance, [Balart & Oosterveen, 2019](#)) argue that female students are more likely to review their answers before handing in a test, less likely to get stuck on a question for a long time, and more likely to verify that all questions have been answered. Similarly, in agreement with the results of several works, our finding confirms the positive association between socioeconomic background and non-cognitive skills. As reported in Section 2.2.2, children from wealthier families have more opportunities to become perseverant and to develop a positive attitude towards learning than their peers from poorer families. Following from this, our results indicate that the gender and socioeconomic gaps in digital skills both decrease after accounting for the influence of non-cognitive variables.

On the other hand, at variance with the existing literature, our estimates suggest that non-cognitive skills work towards widening the native-immigrant gap in digital skills. A possible explanation for this discrepancy is that previous studies on the relationship between immigrant background and non-cognitive skills rely on self-reported proxies for non-cognitive skills, whereas we employ an indicator that considers students' behaviour on the test. Although many immigrant students report high intrinsic learning motivation, there is the possibility that this is essentially the result of the pressure placed on them by public opinion or their parents. They may feel socially obliged to report this in order to avoid negative stereotypes ([Suarez-Orozco et al., 2017](#)). Alternatively, while a positive attitude towards education and higher engagement in learning are generally found to be associated with better academic performance, this relationship may be weaker or non-existent for some groups of students, such as those with an immigrant background — this is often referred as the so-called “engagement-achievement paradox” ([Shernoff & Schmidt, 2008](#)). One reason for it is that, though immigrant parents pass on their academic aspirations and expectations to their children, their ability to actually support and shape the learning of their offspring is limited by a number of obstacles. For instance, due to language barriers and lack of familiarity with the host educational system, they may experience a poor understanding or mistrust of school-related feedback and have poor confidence in their ability to support children's homework ([Yamamoto & Holloway, 2010](#)).

This paper shows that, to be informative, results from large-scale educational assessments in all domains (including digital skills) should account for the fact that test scores are affected by non-cognitive skills that considerably vary across students both within and across countries. A standard strategy to handle this issue should be agreed upon and only results adjusted for differences in non-cognitive skills should be used for cross-country comparisons and accountability purposes. Otherwise, as

shown by this study, there is, for instance, the risk that the importance of gender in explaining digital performance differences across students may be overstated. Where possible, policymakers should base the adoption of relevant education policies on valid, reliable, and comparable findings.

6.1. Limitations

Our study is not free from limitations. First, the country/region/city coverage of the ICILS data is somewhat limited. Extending the number of participants is needed to improve the generalisability of our results.

Second, although, as mentioned in Section 1, cognitive and non-cognitive skills tend to capture different dimensions with separate effects on students' outcomes, one cannot completely rule out the possibility that our proxies for non-cognitive traits pick up the effect of some cognitive capacities. For example, while analysing the impact of item position effects on school test results, Nagy et al. (2018) find that decoding speed, which is a prototypical exemplar of cognitive capacities, is key to maintain a constant level of effort in an exam.

Third, differences between students in terms of digital skills are also likely to be related to differences across countries/regions/cities in terms, for instance, of ICT infrastructure and ICT educational programmes. Unfortunately, in our analysis we are unable to clearly identify these factors and their relative importance. The unobserved heterogeneity that may be present in fixed-effects models can be considered as a "black box" (Hill et al., 2020). Thus, future analyses using the ICILS data could benefit from explicitly modelling the effects of country/region/city-level characteristics on students' CIL test scores.

Furthermore, digital skills are a broad and complex construct, comprising different dimensions and different types of abilities. For example, the ability to digitally edit a photograph and post it on a website requires a different skill-set than, say, writing a macro to automate calculations in a spreadsheet. Our analysis is limited in that it focuses on a generic, broad indicator of ICT competency. Additional research is needed to investigate if the findings of the present study hold also for more specific domains of digital skills.

Finally, our results about the impact of non-cognitive skills on ICT test scores cannot be given a causal interpretation: there are likely to be (additional) observable or unobservable factors affecting both ICT test scores and non-cognitive skills. In addition, our analysis does not shed any light on the mechanisms through which non-cognitive skills may affect ICT test scores.

Declaration of competing interest

None.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.lindif.2022.102254>.

References

- Aesaert, K., & van Braak, J. (2015). Gender and socioeconomic related differences in performance based ict competences. URL: *Computers & Education*, 84, 8–25. <https://doi.org/10.1016/j.compedu.2014.12.017> <http://www.sciencedirect.com/science/article/pii/S0360131515000020>.
- Alivernini, F., Manganelli, S., Cavicchiolo, E., Girelli, L., Biasi, V., & Lucidi, F. (2018). Immigrant background and gender differences in primary students' motivations toward studying. URL *The Journal of Educational Research*, 111, 603–611. <https://doi.org/10.1080/00220671.2017.1349073>. doi:10.1080/00220671.2017.1349073, arXiv:doi:10.1080/00220671.2017.1349073.
- Anghel, B., & Balart, P. (2017). Non-cognitive skills and individual earnings: New evidence from piae. *SERIES*, 8, 417–473.
- Arepattamanni, S., & Freeman, J. G. (2008). Academic achievement, academic self-concept, and academic motivation of immigrant adolescents in the greater toronto area secondary schools. URL: *Journal of Advanced Academics*, 19, 700–743. <https://doi.org/10.4219/jaa-2008-831>. doi:10.4219/jaa-2008-831, arXiv:doi:10.4219/jaa-2008-831.
- Arellano, A., Cámara, N., & Tuesta, D. (2018). Explaining the gender gap in financial literacy: The role of non-cognitive skills. URL: *Economic Notes*, 47, 495–518. <https://doi.org/10.1111/ecn.12113>. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecn.12113 <https://onlinelibrary.wiley.com/doi/abs/10.1111/ecn.12113>.
- Azzolini, D., Bazoli, N., Lievore, I., Schizzerotto, A., & Vergolini, L. (2019). *Beyond achievement: A comparative look into 15-year-olds' school engagement, effort and perseverance in the European Union*. Technical report. Publications Office of the European Union.
- Balart, P., & Oosterveen, M. (2019). Females show more sustained performance during test-taking than males. *Nature Communications*, 10, 3798. <https://doi.org/10.1038/s41467-019-11691-y>. URL: doi:10.1038/s41467-019-11691-y.
- Balart, P., Oosterveen, M., & Webbink, D. (2018). Test scores, noncognitive skills and economic growth. URL: *Economics of Education Review*, 63, 134–153. <https://doi.org/10.1016/j.econedurev.2017.12.004> <https://www.sciencedirect.com/science/article/pii/S0272775717306854>.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.
- Bidjerano, T., & Dai, D. Y. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. URL: *Learning and Individual Differences*, 17, 69–81. <https://doi.org/10.1016/j.lindif.2007.02.001> <https://www.sciencedirect.com/science/article/pii/S104160800700026X>.
- Borghans, L., Meijers, H., & ter Weel, B. (2008). The role of noncognitive skills in explaining cognitive test scores. URL: *Economic Inquiry*, 46, 2–12. <https://doi.org/10.1111/j.1465-7295.2007.00073.x> <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1465-7295.2007.00073.x>.
- Borghans, L., & Schils, T. (2018). *Decomposing Achievement Test Scores into Measures of Cognitive and Noncognitive Skills (December 20, 2018)*. Technical Report. Available at SSRN: <https://doi.org/10.2139/ssrn.3414156>. or <https://ssrn.com/abstract=3414156>.
- Borgonovi, F., & Biecek, P. (2016). An international comparison of students' ability to endure fatigue and maintain motivation during a low-stakes test. URL: *Learning and Individual Differences*, 49, 128–137. <https://doi.org/10.1016/j.lindif.2016.06.001> <https://www.sciencedirect.com/science/article/pii/S1041608016300851>.
- Breakspear, S. (2012). *The Policy Impact of PISA*. URL: <https://doi.org/10.1787/5k9fdqf28-en> <https://www.oecd-ilibrary.org/content/paper/5k9fdqf28-en>.
- Choi, Á., & Jerrim, J. (2016). The use (and misuse) of PISA in guiding policy reform: The case of Spain. URL: *Comparative Education*, 52, 230–245. <https://doi.org/10.1080/03050068.2016.1142739>. doi:10.1080/03050068.2016.1142739, arXiv:doi:10.1080/03050068.2016.1142739.
- Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical methods for comparing regression coefficients between models. URL *American Journal of Sociology*, 100, 1261–1293 <http://www.jstor.org/stable/2782277>.
- Conrad, N., & Patry, M. W. (2012). Conscientiousness and academic performance: A mediational analysis. *International Journal for the Scholarship of Teaching and Learning*, 6. <https://doi.org/10.20429/ijsoil.2012.060108>
- Costa, P. T., Jr., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, 81, 322.
- DeMars, C. E., Bashkov, B. M., & Socha, A. B. (2013). The role of gender in test-taking motivation under low-stakes conditions. *Research & Practice in Assessment*, 8, 69–82.
- Di Chiacchio, C., De Stasio, S., & Fiorilli, C. (2016). Examining how motivation toward science contributes to omitting behaviours in the Italian PISA 2006 sample. URL: *Learning and Individual Differences*, 50, 56–63. <https://doi.org/10.1016/j.lindif.2016.06.025> <https://www.sciencedirect.com/science/article/pii/S104160801630108X>.
- DiPrete, T. A., & Jennings, J. L. (2012). Social and behavioral skills and the gender gap in early educational achievement. URL: *Social Science Research*, 41, 1–15. <https://doi.org/10.1016/j.ssresearch.2011.09.001> <https://www.sciencedirect.com/science/article/pii/S0049089X11001402>.
- Dobbins, M., & Martens, K. (2012). Towards an Education Approach à La Finlandaise? French education policy after PISA. URL *Journal of Education Policy*, 27, 23–43. <https://doi.org/10.1080/02680939.2011.622413>. doi:10.1080/00220671.2017.1349073, arXiv:doi:10.1080/00220671.2017.1349073.
- Drechsler, J. (2020). Multiple imputation. URL: In P. Atkinson, S. Delamont, A. Cernat, J. W. Sakshaug, & R. A. Williams (Eds.), *SAGE Research Methods Foundations*. London: Sage Publications. <https://doi.org/10.4135/9781526421036885886>.
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92, 1087.
- Duckworth, A. L., Tsukayama, E., & May, H. (2010). Establishing causality using longitudinal hierarchical linear modeling: An illustration predicting achievement from self-control. URL: *Social Psychological and Personality Science*, 1, 311–317. <https://doi.org/10.1177/1948550609359707>. doi:10.1177/1948550609359707, arXiv:doi:10.1177/1948550609359707. PMID: 20976121.
- Enders, C. K., & Tofghi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, 12, 121–138.
- Farkas, G. (2003). Cognitive skills and noncognitive traits and behaviors in stratification processes. URL: *Annual Review of Sociology*, 29, 541–562 <http://www.jstor.org/stable/30036979>.
- Farsides, T., & Woodfield, R. (2003). Individual differences and undergraduate academic success: The roles of personality, intelligence, and application. URL: *Personality and Individual Differences*, 34, 1225–1243. [https://doi.org/10.1016/S0191-8869\(02\)00111-3](https://doi.org/10.1016/S0191-8869(02)00111-3) <https://www.sciencedirect.com/science/article/pii/S0191886902001113>.

- Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin*, 116, 429.
- Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Duckworth, D. (2020). *IEA international computer and information literacy 2018: Technical report*. URL: Amsterdam: International Association for the Evaluation of Educational Achievement. doi:10.1080/03050068.2016.1142739, arXiv:doi:10.1080/03050068.2016.1142739 <https://www.iea.nl/index.php/publications/technical-reports/icils-2018-technical-report>.
- Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Duckworth, D. (2020). *Preparing for life in a digital world. IEA international computer and information literacy study 2018 international report*. Springer International Publishing.
- Fulgini, A. J., & Tseng, V. (1999). Family obligations and the achievement motivation of children from immigrant and American-born families. *Advances in Motivation and Achievement*, 159–184.
- Gebhardt, E., Thomson, S., Ainley, J., & Hillman, K. (2019). In *What have we learned about gender differences in ICT?* (pp. 69–73). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-26203-7_6. URL: doi:10.1007/978-3-030-26203-7_6.
- Gnambs, T. (2021). The development of gender differences in information and communication technology (ICT) literacy in middle adolescence. URL: *Computers in Human Behavior*, 114, Article 106533. <https://doi.org/10.1016/j.chb.2020.106533> <http://www.sciencedirect.com/science/article/pii/S0747563220302855>.
- Gneezy, U., List, J. A., Livingston, J. A., Qin, X., Sadoff, S., & Xu, Y. (2019). Measuring success in education: The role of effort on the test itself. URL: *American Economic Review: Insights*, 1, 291–308. <https://doi.org/10.1257/aeri.20180633> <https://www.aeaweb.org/articles?id=10.1257/aeri.20180633>.
- Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (1994). Role of parental motivational practices in children's academic intrinsic motivation and achievement. *Journal of Educational Psychology*, 86, 104.
- Grundke, R., Marcolin, L., Nguyen, T. L. B., & Squicciarini, M. (2018). *Which skills for the digital era?* URL: <https://doi.org/10.1787/9a9479b5-en> <https://www.oecd-ilibrary.org/content/paper/9a9479b5-en>.
- Guggens, J. (2021). On the predictors of computational thinking and its growth at the high-school level. *Computers & Education*, 161, Article 104060. <https://doi.org/10.1016/j.compedu.2020.104060>. URL: doi:10.1038/s41467-019-11691-y <https://www.sciencedirect.com/science/article/pii/S036013152030258X>.
- Gujarati, D. N. (2004). *Basic econometrics* (4th ed.). The McGraw Hill Companies.
- Gutman, L. M., & Schoon, I. (2013). *The impact of non-cognitive skills on outcomes for young people. A literature review*. London: Education Endowment Foundation.
- Hagger, M. S., & Hamilton, K. (2019). Grit and self-discipline as predictors of effort and academic attainment. URL: doi:10.1177/0081175019852763 *British Journal of Educational Psychology*, 89, 324–342. <https://doi.org/10.1111/bjep.12241>. arXiv: doi:10.1177/0081175019852763 <https://bpspsychub.onlinelibrary.wiley.com/doi/abs/10.1111/bjep.12241>.
- Hatlevik, O. E., Guðmundsdóttir, G. B., & Loi, M. (2015). Digital diversity among upper secondary students: A multilevel analysis of the relationship between cultural capital, self-efficacy, strategic use of information and digital competence. URL: *Computers & Education*, 81, 345–353. <https://doi.org/10.1016/j.compedu.2014.10.019> <http://www.sciencedirect.com/science/article/pii/S0360131514002395>.
- Hatlevik, O. E., Ottestad, G., & Throndsen, I. (2015). Predictors of digital competence in 7th grade: A multilevel analysis. URL: *Journal of Computer Assisted Learning*, 31, 220–231. <https://doi.org/10.1111/jcal.12065> <https://onlinelibrary.wiley.com/doi/abs/10.1111/jcal.12065>.
- Hatlevik, O. E., Throndsen, I., Loi, M., & Gudmundsdottir, G. B. (2018). Students' ICT self-efficacy and computer and information literacy: Determinants and relationships. URL: *Computers & Education*, 118, 107–119. <https://doi.org/10.1016/j.compedu.2017.11.011> <http://www.sciencedirect.com/science/article/pii/S0360131517302555>.
- Hawthorne, K. A., Bol, L., Pribesh, S., & Suh, Y. (2015). Effects of motivational prompts on motivation, effort, and performance on a low-stakes standardized test. *Research & Practice in Assessment*, 10, 30–38.
- Heckman, J. J., Jagelka, T., & Kautz, T. D. (2019). *Some contributions of economics to the study of personality. Working paper 26459*. URL: National Bureau of Economic Research. <https://doi.org/10.3386/w26459> <http://www.nber.org/papers/w26459>.
- Hedengren, D., & Stratmann, T. (2012). *The dog that didn't bark: What item nonresponse shows about cognitive and non-cognitive ability*. Available at SSRN 2194373.
- Hill, T. D., Davis, A. P., Roos, J. M., & French, M. T. (2020). Limitations of fixed-effects models for panel data. URL: doi:10.1177/0731121419863785 *Sociological Perspectives*, 63, 357–369. <https://doi.org/10.1177/0731121419863785>. arXiv:doi: 10.1177/0731121419863785.
- Hitt, C., Trivitt, J., & Cheng, A. (2016). When you say nothing at all: The predictive power of student effort on surveys. URL: *Economics of Education Review*, 52, 105–119. <https://doi.org/10.1016/j.econedurev.2016.02.001> <http://www.sciencedirect.com/science/article/pii/S0272775716300541>.
- Jæger, M. M. (2011). Does cultural capital really affect academic achievement? New evidence from combined sibling and panel data. URL: *Sociology of Education*, 84, 281–298. <https://doi.org/10.1177/0038040711417010>. doi:10.1177/0038040711417010, arXiv:doi:10.1177/0038040711417010.
- Juhaňák, L., Zounek, J., Záleská, K., Bárta, O., & Vlčková, K. (2019). The relationship between the age at first computer use and students' perceived competence and autonomy in ICT usage: A mediation analysis. URL: *Computers & Education*, 141, Article 103614. <https://doi.org/10.1016/j.compedu.2019.103614> <https://www.sciencedirect.com/science/article/pii/S0360131519301678>.
- Kattan, R. B. (2017). Non-cognitive skills: What are they and why should we care?. URL: <https://blogs.worldbank.org/education/non-cognitive-skills-what-are-they-and-why-should-we-care>.
- Khine, M. S., & Areepattamannil, S. (2016). *Non-cognitive skills and factors in educational attainment. volume 9 of contemporary approaches to research in learning innovations*. URL: Rotterdam: Brill.
- Khorramdel, L., von Davier, M., Gonzalez, E., & Yamamoto, K. (2020). URL: In *Plausible values: Principles of item response theory and multiple imputations* (pp. 27–47). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-47515-4_3. doi:10.1080/00220671.2017.1349073, arXiv:doi:10.1080/00220671.2017.1349073.
- Kormos, J., & Csizér, K. (2014). The interaction of motivation, self-regulatory strategies, and autonomous learning behavior in different learner groups. URL: *TESOL Quarterly*, 48, 275–299. <https://doi.org/10.1002/tesq.129> <https://onlinelibrary.wiley.com/doi/abs/10.1002/tesq.129>.
- Kormos, J., & Kiddle, T. (2013). The role of socio-economic factors in motivation to learn English as a foreign language: The case of Chile. URL: *System*, 41, 399–412. <https://doi.org/10.1016/j.system.2013.03.006> <https://www.sciencedirect.com/science/article/pii/S0346251X13000523>.
- Lechner, C. M., Anger, S., & Rammstedt, B. (2019). Socio-emotional skills in education and beyond: Recent evidence and future research avenues. In R. Becker (Ed.), *Research handbook on the sociology of education* (pp. 427–453). Edward Elgar Publishing.
- Lechner, C. M., Bhaktha, N., Groskurth, K., & Bluemke, M. (2021). Why ability point estimates can be pointless: A primer on using skill measures from large-scale assessments in secondary analyses. *Measurement Instruments for the Social Sciences*, 3, 1–16. <https://doi.org/10.1186/s42409-020-00020-5>.
- Lee, J. (2020). *Non-cognitive characteristics and academic achievement in Southeast Asian countries based on PISA 2009, 2012, and 2015*. URL: doi:10.1177/0081175019852763. <https://doi.org/10.1787/c3626e2f-en>. arXiv:doi:10.1177/0081175019852763 <https://www.oecd-ilibrary.org/content/paper/c3626e2f-en>.
- Liu, A. (2019). Can non-cognitive skills compensate for background disadvantage? – The moderation of non-cognitive skills on family socioeconomic status and achievement during early childhood and early adolescence. URL: *Social Science Research*, 83, Article 102306. <https://doi.org/10.1016/j.ssresearch.2019.04.019> <https://www.sciencedirect.com/science/article/pii/S0049089X18304939>.
- Lumley, T. (2010). *Complex surveys: A guide to analysis using R: A guide to analysis using R*. John Wiley and Sons.
- Lumley, T. (2020). *Survey: Analysis of complex survey samples. R package version 4.0*.
- Lundahl, C., & Serder, M. (2020). Is pisa more important to school reforms than educational research? The selective use of authoritative references in media and in parliamentary debates. URL: doi:10.1080/20020317.2020.1831306 *Nordic Journal of Studies in Educational Policy*, 6, 193–206. <https://doi.org/10.1080/20020317.2020.1831306>. arXiv:doi:10.1080/20020317.2020.1831306.
- Meyer, J., Fleckenstein, J., Retelsdorf, J., & Köller, J. (2019). The relationship of personality traits and different measures of domain-specific achievement in upper secondary education. *Learning and Individual Differences*, 69, 45–59. <https://doi.org/10.1016/j.lindif.2018.11.005>.
- Mikheeva, E., & Meyer, S. (2020). *IEA international computer and information literacy 2018: User guide for the international database*. URL: The International Association for the Evaluation of Educational Achievement (IEA).
- Mize, T. D., Doan, L., & Long, J. S. (2019). A general framework for comparing predictions and marginal effects across models. URL: doi:10.1177/0081175019852763 *Sociological Methodology*, 49, 152–189. <https://doi.org/10.1177/0081175019852763>. arXiv:doi:10.1177/0081175019852763.
- Morandini, M. C., Thum-Thysen, A., & Vandeplass, A. (2020). *Facing the digital transformation: Are digital skills enough? European Economy - economic briefs 054*. Directorate General Economic and Financial Affairs (DG ECFIN). URL: European Commission <https://ideas.repec.org/p/euf/ecobri/054.html>.
- Munger, K., Gopal, I., Nagler, J., & Tucker, J. A. (2021). Accessibility and generalizability: Are social media effects moderated by age or digital literacy?. URL: doi:10.1177/20531680211016968 *Research & Politics*, 8, Article 20531680211016968. <https://doi.org/10.1177/20531680211016968>. arXiv:doi: 10.1177/20531680211016968.
- Nagy, G., Nagengast, B., Becker, M., Rose, N., & Frey, A. (2018). Item position effects in a reading comprehension test: An irt study of individual differences and individual correlates. *Psychological Test and Assessment Modeling*, 60, 165–187.
- Palacios-Abad, A. (2021). Strive to succeed? The role of persistence in the process of educational attainment. URL: *American Behavioral Scientist*. <https://doi.org/10.1177/0002764221996758>.
- Punter, R. A., Meelissen, M. R., & Glas, C. A. (2017). Gender differences in computer and information literacy: An exploration of the performances of girls and boys in ICLIS 2013. URL: doi:10.1177/1474904116672468 *European Educational Research Journal*, 16, 762–780. <https://doi.org/10.1177/1474904116672468>. arXiv:doi:10.1177/1474904116672468.
- Rozon, O. B. (2013). *Standardized tests: What factors affects how a student does on these exams? Technical report*. URL: University of Puget Sound. doi:10.1177/1948550609359707, arXiv:doi:10.1177/1948550609359707. pMID: 20976121 https://www.pugetsound.edu/sites/default/files/file/rozonstandardizedtestswhatfactorsaffecthowstudent_0.pdf.
- Scherer, R., & Siddiq, F. (2019). The relation between students' socioeconomic status and ICT literacy: Findings from a meta-analysis. URL: *Computers & Education*, 138, 13–32. <https://doi.org/10.1016/j.compedu.2019.04.011> <http://www.sciencedirect.com/science/article/pii/S0360131519300983>.
- Schleicher, A. (2006). Where immigrant students succeed: A comparative review of performance and engagement in PISA 2003. URL: doi:10.1080/

- 14675980601063900 *Intercultural Education*, 17, 507–516. <https://doi.org/10.1080/14675980601063900>. arXiv:doi:10.1080/14675980601063900.
- Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in big five personality traits across 55 cultures. *Journal of Personality and Social Psychology*, 94, 168.
- Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating Cohen's, a measure of local effect size, from proc mixed. arXiv: *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00111> <https://www.frontiersin.org/articles/10.3389/fpsyg.2012.00111/full>.
- Shernoff, D. J., & Schmidt, J. A. (2008). Further evidence of an engagement–achievement paradox among U.S. high school students. *Journal of Youth and Adolescence*, 37, 564–580. <https://doi.org/10.1007/s10964-007-9241-z>. URL: doi:10.1007/s10964-007-9241-z.
- Siddiq, F., & Scherer, R. (2019). Is there a gender gap? A meta-analysis of the gender differences in students' ICT literacy. URL: *Educational Research Review*, 27, 205–217. <https://doi.org/10.1016/j.edurev.2019.03.007> <https://www.sciencedirect.com/science/article/pii/S1747938X18302902>.
- Soland, J., Zammaro, G., Cheng, A., & Hitt, C. (2019). Identifying naturally occurring direct assessments of social-emotional competencies: The promise and limitations of survey and assessment disengagement metadata. URL: *Educational Researcher*, 48, 466–478. <https://doi.org/10.3102/0013189X19861356>. doi:10.4219/jaa-2008-831, arXiv:doi:10.4219/jaa-2008-831.
- St-Hilaire, A. (2002). The social adaptation of children of Mexican immigrants: Educational aspirations beyond junior high school. URL: *Social Science Quarterly*, 83, 1026–1043. <https://doi.org/10.1111/1540-6237.00131>. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/1540-6237.00131 <https://onlinelibrary.wiley.com/doi/abs/10.1111/1540-6237.00131>.
- Strayhorn, T. L. (2014). What role does grit play in the academic success of black male collegians at predominantly white institutions? *Journal of African American Studies*, 18, 1–10. <https://doi.org/10.1007/s12111-012-9243-0>
- Suarez-Orozco, C., Suarez-Orozco, M., & Strom, A. (2017). Immigrant students are internalizing stereotypes. Educators can help. URL: <https://www.edweek.org/leadership/opinion-immigrant-students-are-internalizing-stereotypes-educators-can-help/2017/09>.
- Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. URL *Journal of Personality*, 72, 271–324. <https://doi.org/10.1111/j.0022-3506.2004.00263.x>. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.0022-3506.2004.00263.x <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0022-3506.2004.00263.x>.
- Thorsen, C., Yang Hansen, K., & Johansson, S. (2021). The mechanisms of interest and perseverance in predicting achievement among academically resilient and non-resilient students: Evidence from Swedish longitudinal data. URL: *British Journal of Educational Psychology*, Article e12431. <https://doi.org/10.1111/bjep.12431>. arXiv:https://bpspsychub.onlinelibrary.wiley.com/doi/pdf/10.1111/bjep.12431 <https://bpspsychub.onlinelibrary.wiley.com/doi/abs/10.1111/bjep.12431>.
- Tseng, V. (2002). *Family as a context for immigrant adaptation: Family interdependence adjustment and course of study among youths from immigrant and United States born families*. URL: doi:10.1177/1948550609359707, arXiv:doi:10.1177/1948550609359707. PMID: 20976121.
- Urduan, T. (2004). Predictors of academic self-handicapping and achievement: Examining achievement goals, classroom goal structures, and culture. *Journal of Educational Psychology*, 96, 251.
- Usher, A., & Kober, N. (2012). 4. *What roles do parent involvement, family background, and culture play in student motivation?* URL: Center on Education Policy.
- Wang, W., Dong, Y., Liu, X., Bai, Y., & Zhang, L. (2020). The effect of parents' education on the academic and non-cognitive outcomes of their children: Evidence from China. URL: *Children and Youth Services Review*, 117, Article 105307. <https://doi.org/10.1016/j.childyouth.2020.105307> <https://www.sciencedirect.com/science/article/pii/S0190740920312093>.
- Wehner, C., & Schils, T. (2021). Who are the low educational achievers? An analysis in relation to gender, emotional stability and conscientiousness. URL: doi:10.1080/00036846.2021.1922592 *Applied Economics*, 53, 5354–5368. <https://doi.org/10.1080/00036846.2021.1922592>. arXiv:doi:10.1080/00036846.2021.1922592.
- Xu, K. M., Cunha-Harvey, A. R., King, R. B., de Koning, B. B., Paas, F., Baars, M., Zhang, J., & de Groot, R. (2021). A cross-cultural investigation on perseverance, self-regulated learning, motivation, and achievement. URL: doi:10.1080/03057925.2021.1922270 *Compare: A Journal of Comparative and International Education*, 0, 1–19. <https://doi.org/10.1080/03057925.2021.1922270>. arXiv:doi:10.1080/03057925.2021.1922270.
- Yamamoto, Y., & Holloway, S. D. (2010). Parental expectations and children's academic performance in sociocultural context. *Educational Psychology Review*, 22, 189–214. <https://doi.org/10.1007/s10648-010-9121-z>
- Yunus, M., & Ali, W. Z. W. (2009). Motivation in the learning of mathematics. *European Journal of Social Sciences*, 7, 93–101.
- Zammaro, G., Hitt, C., & Mendez, I. (2019). When students don't care: Reexamining international differences in achievement and student effort *Journal of Human Capital*, 13, 519–552. <https://doi.org/10.1086/705799>. URL: doi:10.1007/978-3-030-26203-7_6.
- Zammaro, G., Nichols, M., Duckworth, A. L., & D'Mello, S. K. (2020). Validation of survey effort measures of grit and self-control in a sample of high school students. *PLOS ONE*, 15, 1–16. <https://doi.org/10.1371/journal.pone.0235396>. URL: 10.1371/journal.pone.0235396.
- Zigler, E., & Butterfield, E. C. (1968). Motivational aspects of changes in iq test performance of culturally deprived nursery school children. URL *Child Development*, 39, 1–14 <http://www.jstor.org/stable/1127353>.
- Zilberberg, A., Finney, S. J., Marsh, K. R., & Anderson, R. D. (2014). The role of students' attitudes and test-taking motivation on the validity of college institutional accountability tests: A path analytic model. *International Journal of Testing*, 14, 360–384. <https://doi.org/10.1080/15305058.2014.928301>. URL: doi:10.1007/s10964-007-9241-z.