Late-childhood foundational cognitive skills predict educational outcomes through adolescence and into young adulthood: evidence from Ethiopia and Peru

19th April 2024

Jennifer Lopez, [a] Jere Behrman, [b] Santiago Cueto, [c] Marta Favara, [d] Alan Sánchez [e]

[a] *Grupo de Análisis para el Desarrollo (GRADE), Lima, Peru. Corresponding author (jelopez@grade.org.pe)*

[b] *Department of Economics, University of Pennsylvania, Philadelphia, USA. (jbehrman@econ.upenn.edu)*

[c] *Grupo de Análisis para el Desarrollo (GRADE) and Pontificia Universidad Catolica del Peru Lima, Peru (scueto@grade.org.pe)*

[d] *Oxford Department of International Development, University of Oxford, Oxford, UK. (marta.favara@qeh.ox.ac.uk)*

[e] *Grupo de Análisis para el Desarrollo (GRADE), Lima, Peru and Oxford Department of International Development, University of Oxford, Oxford, UK. (asanchez@grade.org.pe)*

Ethical Approval: The Institutional Review Board of the University of Pennsylvania approved this project (protocol #834313).

Acknowledgments: This study was funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development project entitled "Foundational cognitive skills in developing countries: early-life nutritional, climatic and policy determinants and impacts on adolescent education, socio-emotional competencies and risky behaviors." under grant number NICHD R21 HD097576. The funder had no role in the design, interpretation or writing-up of the study or in the decision to submit the study for consideration for publication. Thanks also to the UK's Foreign, Commonwealth and Development Office (FCDO) for funding Young Lives at Work and enabling this research and to Old Dart Foundation for providing additional funding for this study. We are grateful to Margaret Sheridan, Richard Freund, Douglas Scott, Annina Hittmeyer and Nicolas Pazos for providing very useful comments on earlier drafts.

Abstract

We estimate associations between foundational cognitive skills (inhibitory control, working memory, long-term memory, and implicit learning) measured at age 12 and educational outcomes measured at ages 15 and 19-20 in Ethiopia and Peru, using the Young Lives data. The estimates adjust for rich sets of controls and include measurements of children's baseline abilities. For a subset of the outcomes, we exploit within-household variation. Working memory and long-term memory are consistently and positively associated with subsequent domain-specific cognitive achievement tests (measuring specifically numeracy, vocabulary and literacy achievement) in both countries, university enrolment in Peru (long-term memory) and lower secondary-school completion in Ethiopia (working memory). Inhibitory control predicts subsequent math-test scores in both countries, grade attainment (Ethiopia), and university enrolment (Peru). Value-added estimates show that these skills play roles during adolescence, with the memory-related skills predicting higher domain-specific test scores (Peru and Ethiopia) and grade attainment (Ethiopia), while inhibitory control has associations with math (both countries). These results provide additional evidence to justify the importance of promoting investments in cognitive skills throughout childhood and adolescence, and elucidate how such investments impact educational achievements.

Keywords: human capital; cognitive skills; education; executive function; Ethiopia; Peru

JEL codes: I25, I24, I23

1. Introduction

The promotion of learning opportunities for all is an important challenge for developing countries, as recognized by Sustainable Development Goal N° 4 (United Nations, 2015). Most studies carried out in low-and-middle-income countries (LMICs) about the impact of educational inputs have focused on school resources and processes, including, for example, the roles of teachers (Glewwe et al., 2011; McEwan, 2015). While cognitive skills (CSs) are thought to play important roles in educational attainment, which in turn determines how well children can succeed in economic outcomes later in life (Hanushek, 2009; Hanushek & Woessmann, 2008; Reynolds et al., 2010), data about CSs are limited in LMICs. Furthermore, CSs are not fully captured by domain-specific achievement tests, commonly used to approximate cognitive skills. CSs include working and long-term memory, planning and problem solving, attention control, processing speed, spatial and numerical processing, and social cognition, among other domains. Recent theoretical and empirical developments have suggested the importance of a group of these skills under the umbrella of what is called executive functions (EFs). While there is not unanimous agreement among researchers on which skills form EFs, most definitions include working memory, inhibition control, and attention flexibility or control (Carlson, 2005; Garon et al., 2008). EFs permit regulation of thought and action (Miyake & Friedman, 2012), which allows children to adapt and perform well in classroom environments. EFs are also malleable throughout adolescence due to their associations with connections in the prefrontal cortex, an area that undergoes changes into young adulthood (Benes, 2001).

Although CSs, including EFs, are essential to learning and thus expected to be linked to a variety of learning outcomes over the life cycle, there is limited evidence about which CSs predict different domain-specific learning outcomes in LMICs. We contribute to fill this gap by analysing longitudinal data collected in late childhood in two LMICs using a novel computerized method to measure inhibitory control, working memory, long-term memory, and implicit learning. We refer to these four skills as Foundational Cognitive Skills (FCSs) throughout this paper—and use the term CSs to refer to research of other authors measuring only some of these skills, as well as others, including domain-specific ones. Specifically, the purpose of this paper is to investigate the predictive value of FCSs on educational achievements among adolescents and young adults growing-up in LMIC settings in Ethiopia and Peru. To do this, we assess the relationships between FCSs in late childhood at age 12 with domain-specific cognitive test scores and grade attainment in adolescence at age 15, and with lower-secondary-school completion in Ethiopia and university enrolment in Peru in young adulthood at ages 19 and 20, respectively. We use longitudinal data from the Young Lives (YL) study; specifically, we use information from a sample of children tracked from age 1 (in 2002) to age 20 (in 2021) in both Ethiopia and Peru. FCSs are measured by the Rapid Assessment of Cognitive and Emotional Regulation (RACER) (Behrman et al., 2022; Hamoudi & Sheridan, 2015), which is a tablet-based assessment application administered in the Ethiopian and Peruvian YL samples in 2013. Even though the two samples are not representative of the general populations of these countries, they capture a wide spectrum of living standard conditions observed in both countries (Escobal & Flores, 2008; Outes-Leon & Sánchez, 2008).

Our study offers three main contributions. First, we use unique data on FCSs. To our knowledge, this is the first longitudinal data set containing information about domain-specific cognitive skills and education attainments in LMICs. Unlike achievement tests, FCSs measures are not domain-specific, and should, therefore, be relatively free of bias due to the language of implementation or cultural differences. Second, we investigate the potential role of FCSs in improving educational outcomes during adolescence. Most research has focused on the development of CSs (see, for instance, Garon et al., 2008, Best & Miller, 2010) and its relationship with achievement (e.g., Cain et al., 2004) during the preschool period. However, there is evidence that the latent organization of CSs components changes from the preschool ages to adulthood (Lee et al., 2013; Miller et al., 2012; Miyake et al., 2000; Monette et al., 2015; Usai et al., 2014). This means that some components of CSs, including some of those related to EFs, can be altered even after the preschool period, which is to say that those CSs can be malleable even after the first years of life. Therefore, policies aimed at improving learning may be effective and serve as remediation even for older children (Freund et al., 2023; Scott at al., 2022). Third, our data are home-based, not school-based, so our analyses are not confounded by unobserved factors that determine school enrolment and attendance, which is particularly important for adolescence in the contexts studied because in these contexts many adolescents do not attend school regularly or at all.

Our main findings are as follows. First, working memory and long-term memory are consistently and positively associated with subsequent domain-specific cognitive achievement tests in both countries, university enrolment in Peru (long-term memory) and lower secondary-school completion in Ethiopia (working memory). Second, inhibitory control predicts subsequent math test scores in both countries, grade attainment in Ethiopia, and university enrolment in Peru. Third, value-added estimates show that FCSs at age 12 predict educational outcomes at age 15 conditional on those outcomes at age 12, which adds to the literature about the importance of CSs, since it suggests that CSs matter to explain changes in educational outcomes during adolescence.

The remainder of this paper discusses key findings from the main literature about CSs and their relationships with academic achievements (Section 2); describes the YL study samples in Ethiopia and Peru, and the data obtained through RACER (Section 3); presents our estimation strategy and the main results (Sections 4 and 5, respectively); reports the robustness checks performed (Section 6), and concludes with a discussion of the main results (Section 7).

2. Literature review

There has been much research on the relationships between CSs and domain-specific test scores. Overall, the existing evidence suggests that CSs are positively associated with domain-specific cognitive achievements, specifically math, vocabulary and literacy skills. For instance, Cain et al. (2004) find that working memory predicts reading comprehension for a longitudinal sample of children in England evaluated at ages eight, nine and 11. Similarly, Christopher et al. (2012) find that working memory is positively correlated with reading-comprehension and also with processing speed, for a sample of US children, ages eight to 16. McClelland´s investigation (2000) of associations between behavioural regulation, and literacy, vocabulary and math test scores concludes that behavioural aspects of self-regulation are important for the development of early domain-specific skills and school success in US children ages four and five. In this case, the main component of behavioural regulation is inhibitory control. Consistent with this result, Merkley et al. (2016) find that inhibitory control is correlated with performance on a math test for a sample of UK children, ages three to five.

However, most of the existing research on CSs development and its relationship to educational outcomes focuses on the preschool period, while there are only a few studies that assess the relationship between CSs and education attainment and performance in late childhood and adolescence (Riggs et al., 2015, Friedman et al., 2016; Friedman et al, 2011; Friedman et al., 2008, Huizinga et al., 2006; McAuley & White, 2011). In addition, there is evidence of the predictive power of CSs on outcomes during adulthood (final education and occupational attainment), as reviewed by Roberts et al. (2007).

Notably, most of the studies investigate the relationship between CSs and educational attainment in the context of high-income countries (HICs, predominantly the United States and the United Kingdom), whereas there are only a few studies using data from LMICs (Christopher et al., 2012; Evans & Popova, 2016). For HICs, Carneiro and Heckman (2002) show that long-term factors associated with family income (and parental education) that promote skills are more important than short-term financial constraints to explain the differences in higher-education enrolment in the US. Put differently, higher-income households can invest more in CSs early in life, as shown by Jerrim & Vignoles (2015), using data from England, Canada, Australia and the US. An environment that can provide the means for children to improve their CSs is more often found in richer households that are more prominent in HICs, in comparison with poorer households that lack the resources to stimulate cognitive growth that are more common in LMICs. In many cases, poorer households also have parents with lower levels of education, which can also limit their children´s development (Stipek & Ryan, 1997; Griffin & Morrison, 1997). Additionally, richer households can also provide better schooling, which can further develop children´s CSs, since it is at school where children continually repeat tasks that develop basic abilities such as working and long-term memory, and attention (Blair, 2002).[i] For LMICs, Castro et al. (2016) find that financial constraints impose a greater deterrent than skills and other family and educational background variables for enrolment in higher education in Peru. Nonetheless, they also find that CSs (numeracy and problem-solving capacity, working memory, verbal fluency, and receptive language) are also significantly associated with enrolment in universities.

The relationships between EFs and early numeracy and early literacy have been widely studied (Blair & Razza, 2007; Bull et al., 2008; Lan et al., 2011; Dilworth-Bart, 2012; Schmitt et al., 2017). Since EFs are a construct, different authors consider different CSs for their empirical representation of EFs. Most of the studies find a positive correlation between EFs and both math and language skills, but this evidence comes mainly from HICs. For example, Blair and Razza (2007) assess the relationships between inhibitory control and attention shifting, and academic abilities (math and reading comprehension) in preschool and kindergarten among US children and find that inhibitory control has strong correlations with both math and reading skills. Likewise, Bull et al. (2008) consider inhibition, shifting and updating as part of EFs, and treat working memory as a separate skill. They conclude that higher levels of EFs are associated with advantages in math and literacy skills in a sample of preschool children from Scotland. Additionally, Lan et al. (2011) and Dilworth-Bart (2012) find that EFs (inhibitory control, working memory and attentional control) is a significant predictor of both mathematics and literacy achievement, using data from a sample of pre-school-aged children in China and the US. The existing literature on EFs using data from LMICs is quite scarce. None of the existing studies examine the relationships between EFs and educational achievement. Instead, they investigate the structure and composition of EFs (Rowe et al., 2021; Wray et al., 2020), compare EFs measures in different contexts (Cheng et al., 2019; Ford et al., 2019; Rowe et al., 2021; Willoughby et al., 2019) or test and validate instruments to measure CSs, like RACER and the OCS-EF instrument from the University of Oxford (Rowe et al., 2021; Yuan et al., 2022).

3. Data and descriptive statistics

3.1 The Young Lives study

YL has been tracking two cohorts of children in Ethiopia, India, Peru and Vietnam: a younger cohort born in 2001-2, and an older cohort born in 1994-5. Our analysis focuses on the younger cohort in Ethiopia and Peru as the RACER data are available only for them.

The initial 2002 survey collected information on 2,052 and 1,999 participants (henceforth, index children), respectively in Ethiopia and Peru, visited in their homes for the first time at ~ age one year. These samples were selected through a two-stage procedure. First, in both countries, 20 clusters were selected (districts in Peru, woredas in Ethiopia). Then, approximately 100 children and their families were randomly selected in each cluster. The selection of clusters aimed at capturing the diversity of each country in terms of geography, climate, ethnicity, and living standards, while also over-representing poor households.[ii] Although the samples were not designed to be nationally representative, comparisons of the distributions of the YL samples with those of Demographic Health Survey (DHS) samples for both countries show that the YL samples capture the diversity of living standards of each country (Escobal & Flores, 2008; Outes-León & Sánchez, 2008).

After the first visit in 2002, four further in-person rounds of data collection took place in 2006, 2009, 2013 and 2016 (Rounds 2-5). During each of these visits, rich information at ~ ages five, eight, 12 and 15, respectively, on household socio-demographic characteristics, children's health and nutrition, educational attainment, achievement test scores, among others, was collected for the index children chosen in the first visit. In addition, a phone survey was administered in 2020-2021 through five phone calls (Call 1-3 in 2020 and Call 4-5 in 2021) to gather information about the impacts of the COVID-19 pandemic on the index children's education, employment, and wellbeing when they were 19-20 years old. For our analysis we use information collected in Round 4, Round 5 and Call 2-5 of the phone survey, when the younger cohort was age 12, 15, 19-20 years old, respectively. Information from Rounds 1-2 is also used to construct some of our control variables.

At the time of the last in-person visit (Round 5), attrition was relatively low compared to other longitudinal studies: 5.4% in Ethiopia and 8.2% in Peru excluding deaths, which represent annual attrition rates of 0.4% and 0.6%, respectively (see Sánchez and Escobal, 2020). The attrition rate of the phone survey was higher due to the remote nature of the phone survey and the impossibility to contact those with outdated contact information. Furthermore, in the case of Ethiopia, the civil conflict that started in late 2020 in some parts of the country led to the decision to not collect data in the Tigray region and in some clusters in nearby regions. As a result, the attrition rate for

Ethiopia increased substantially between late 2020 and 2021. For this reason, for our analysis we use data from Call 2 for Ethiopia (administered between August and October 2020, when the entire sample was contacted) and Call 5 for Peru (administered between October and December 2021). Attrition rates for Ethiopia and Peru in these calls are 13.0% and 17.9%, respectively (with respect to the original 2001 sample).

YL also has collected data from siblings of the original participants since Round 3 in 2009. In Peru, data were collected for 861 younger siblings, ages two to eight years, of the index children, whereas in Ethiopia information was collected for the 1550 siblings closest in age to the index children, 1001 younger siblings ages three to eight years and 449 older siblings ages eight to 17 years. We use data from the younger siblings from both country samples for estimates that exploit within-household variation.

Descriptive statistics of the analytical sample for both countries are reported in Table 1. The Ethiopian sample is mainly rural, whereas the Peruvian sample is mainly urban, reflecting the state of development and urbanization of each country. Furthermore, more than one third of the index children in Ethiopia come from households where the mother has no formal education, compared to less than one tenth in the Peruvian sample. We use mothers' and children's native tongues as proxies for ethnicity. In Ethiopia the three most common mothers' (children´s) native tongues are Amharic, Tigrinya and Oromifa together accounting for 75.4% (80.7%) of the total sample, while in Peru, 70.3% (85.6%) of the sample speak Spanish as their main language. The remaining 24.6% (19.3%) for Ethiopia and 29.7% (14.4%) for Peru are any other native tongues for both countries (Panel A).

INSERT TABLE 1 HERE

Panel B reports descriptive statistics for the index children. By the time of the last in-person visits in 2016 (Round 5), the average age was 15, and 19-20 by the time of the phone surveys in 2020-2021. The educational systems of Ethiopia and Peru are very different, so we present the grade attainments separate for each country. Assuming on-age enrolment and no grade repetition, by age 15, all index children should have reached the ninth grade in Ethiopia and the tenth grade in Peru. However, only 18.4% and 26.2% of the respondents reached these levels by Round 5, respectively.[iii] The variance of grade attainment is larger in Ethiopia, due to more delayed enrolment and grade repetition. By age 15, 40.0% of index children had reached the eighth and ninth grades in the Ethiopian sample, whereas in the Peruvian sample, 74.9% of the sample had reached the ninth and tenth grades. Differences in schooling attainment between the two countries become more pronounced in the transition to young adulthood. According to information collected through the phone survey in Ethiopia, 45.5% had completed lower-secondary school, and 11.7% had completed upper-secondary school by age 19. Furthermore, only 3.1% had ever been enrolled in higher education and only 2.5% were or had been enrolled in a university.[iv] In Peru, 90.0% had completed secondary school, 58.8% were enrolled or had been enrolled in higher education, and 31.8% had ever been enrolled in a university by age 20.

3.2 Cognitive measures in Young Lives

Domain-specific cognitive achievement has been measured in YL starting from age five. The cognitive assessments administered are: receptive vocabulary, numeracy and reading comprehension. The first is measured using the well-known Peabody Picture Vocabulary Test (PPVT) while the latter are assessed using math and reading comprehension tests designed by the YL educational team (Cueto et al., 2009; Cueto & León, 2012). The PPVT has been administered since age five for the index children and age eight for the siblings while the math and reading comprehension tests were administered to the index children only and since the age eight. There is variation in the language of test administration in each country, especially in Ethiopia (Table 1, Panel C). For our analysis, we use the raw scores of each test, which are standardized to have mean zero and variance one within each country sample of index children (who are all very close to the same age). For the sample that includes both index children and siblings, the raw scores are standardized by age in years within each country given the wider age ranges for the siblings. Test scores are standardized by country because test items are adapted to each country´s context, thus the level of difficulty might not be the same.

The RACER tests of FCS were administered in Round 4, when the children were 11-12 years old. RACER is a self-administered, tablet-based software designed to measure FCSs (Sheridan & Hamoudi, 2015). RACER contains tasks designed to measure inhibitory control, working memory, long-term memory, and implicit learning. Inhibitory control and working memory are considered executive functions, i.e., they are part of a set of skills that defines a persons' capacities to achieve their goals. Specifically, inhibitory control refers to one's ability to stop oneself from exhibiting behaviors that one does not want to exhibit and is related to a person's ability to focus on a single task and suppress distractions. Working memory is the ability to hold in mind and manipulate stimuli that are no longer present in the environment. Long-term memory is the ability to encode and retain new knowledge, and then bring it back as relevant for specific tasks. As mentioned in Sections 1 and 2, there is evidence linking executive functions and long-term memory to educational outcomes, but mainly for preschool-age children from HICs. Finally, implicit learning is the ability to recognize and respond to regularities in the environment even when individuals are not aware of these regularities. Implicit learning is linked to language acquisition in young children, as well as learning skills such as riding bikes or swimming (see Amso & Davidow, 2012).

In each RACER task, we can differentiate between baseline and challenge trials. Both were administrated identically. However, the baseline trial helps to control for other skills, aside from the one being assessed, while the challenge trials yield the performance indicator for the cognitive skill measured through each task. Each of the tasks administered to measure these concepts required two-three minutes of instructions and two-four minutes for administration, for a total time for all FCS of approximately 30 minutes. Between 96 and 97% of the sample children took the tests —for the Peruvian sample this is similar to the proportion that took the cognitive achievement tests (PPVT, math, and reading comprehension) at the same age, whereas in the Ethiopian sample a larger proportion of children took the RACER—compared to the cognitive achievement tests (Table 1, Panel D). RACER was also administered to the siblings. Table 2 reports more details on each of the tasks, including which aspects of performance are measured in each case. In the case of long-term memory, due to its nature, the task is presented as the first one and then repeated at the end. Detailed information about the data collected in Ethiopia and Peru is reported in Behrman et al. (2022).

INSERT TABLE 2 HERE

As mentioned, the challenge trials are used to calculate the performance indicator for each FCS. Inhibitory control is computed as the average of the inverse of the average Euclidean distance and the inverse of the average response time, long-term memory as the percentage of correct choices, working memory as the inverse of the average Euclidean distance, and implicit learning as the inverse of the average response time. For the purpose of the statistical analysis, each FCS outcome is standardized and re-scaled so that a higher score is linked to higher ability. Standardization is by age in years for the pooled sample of index children and their siblings in Ethiopia and Peru. As a reference, Table 3 reports average performance across skills by socio-economic status (proxied by household locations and the tertiles in the distributions of the wealth index of each country), and by area of location for each country sample. On average, Peruvian children scored higher than Ethiopian children (in all tests). As expected, children from urban areas scored higher than their counterparts in rural areas, and children from the top tertile of wealth scored higher than those from the lower tertiles. For more details, see Behrman et al. (2022).

INSERT TABLE 3 HERE

4. Empirical methodology

We start with the following specification:

$$EO_{i,c} = \theta_0 + \gamma FCS_i + \theta_1 BT_i + X_i\theta_2 + HC_i\theta_3 + RC_i\theta_4 + \psi_c + \mu_i \qquad (1)$$

where $EO_{i,c}$ is a generic term that represents the outcome of interest observed at age 15 (highest grade attained; standardized scores in the PPVT, math, and reading comprehension tests) or educational attainment at age 19-20 for child i born in community c. Specifically, given the different educational progressions observed in the two countries, the main outcome of interest is having completed lower-secondary education by age 19 in Ethiopia, since general access to secondary education is limited around the nation, and the probability of ever being enrolled in university by age 20 in Peru, as they are the most selective higher-educational institutions in the country. FCS_i is a scalar that represents the score of the child in the challenge trial used to measure each of the FCSs at age ~12. BT_i is an analogous scalar that controls for the performance of the child in the baseline task. The model also includes basic controls at the child level, denoted as vector X_i (age in months, sex, ethnicity proxied by both child's and child´s mother's native tongue, and the number of years the child attended pre-school); and, socio-economic characteristics of the household, denoted as vector HC_i (household wealth[v], maternal level of schooling, and area of residence). RC_i contains additional controls to account for heterogeneity in the administration of the RACER tasks: whether the child completed the tasks during the morning (or not), and during a weekday (or not). ψ_c are community-of-birth fixed effects. Finally, μ_i are unobserved random factors including random measurement errors in $EO_{i,c}$.

The parameter of primary interest is γ. Our specification allows control for several possible sources of bias. First, the fact that the RACER outcomes were measured at age 12, years before the educational outcomes of interest at ages 15, 19 and 20, deals with potential simultaneity bias. Second, although the performance of the child in the challenge task(s) contains information about their general ability, this is conceptually adjusted for by controlling for the performance of the child in the baseline trial(s). Third, the inclusion of the additional controls at the child and household levels adjusts for the roles that the socio-economic status of the household and parental schooling attainment may play in the determination of a child's performance. Fourth, the roles of geographical location—e.g., the quality of the educational services provided in the district—is incorporated through the inclusion of the community-of-birth fixed effects. In Appendix A we report additional results introducing all FCSs at the same time—this permits us to measure the associations of each of these skills with success at school, conditional on the other FCSs.

While Equation (1) helps us determine the associations between FCSs at age 12 and educational outcomes at age 15, it is possible that coefficient estimates might be picking up the role FCSs had on educational outcomes at earlier ages. To try to account for that, in Equation (2) we present a value-added specification that includes the lagged tests scores measured at age 12 as well as all the other controls previously mentioned. In doing so, we can determine whether FCSs at age 12 predict educational outcomes at age 15 conditional on those outcomes at age 12. This is equivalent to the changes in educational outcomes if ρ is not significantly different from one.[vi]

$$EO_{i,c,t} = \theta_0 + \alpha FCS_i + \rho EO_{i,c,t-1} + \theta_1 BT_i + X_i\theta_2 + HC_i\theta_3 + RC_i\theta_4 + \psi_c + \mu_i \quad (2)$$

The way the RACER outcomes were measured reduces the likelihood of omitted-variable bias due to unobserved characteristics at the child level—for example, general ability is controlled for by performance in the baseline tasks. However, results in equation (1) might still be afflicted by omitted-variable bias at the household level. For instance, parental preferences for investing in education might simultaneously explain a child's performance in RACER at age 12 and at school a few years later at age 15, or their probabilities to have completed school or been enrolled in higher education during young adulthood. Parental abilities might play a similar role. These possibilities may be only partially taken into account by controlling for parental schooling and household wealth. To deal with this source of possible omitted-variable bias, additional results are reported using household fixed-effects versions of the main model, taking advantage of the fact that RACER was also administered to the (immediate) younger sibling of the index child in Peru, and the sibling closest in age of the index child in Ethiopia. The older siblings in Ethiopia are excluded from the analysis since their age range is very wide and some of them were already in their early adulthoods when taking RACER. The household fixed-effects specification is as follows:

$$EO_{i,k} = \theta_0 + \gamma FCS_{i,k} + \theta_1 BT_{i,k} + X_{i,k}\theta_2 + HC_k\theta_3 + RC_{i,k}\theta_4 + \psi_k + \mu_{i,k} \quad (3)$$

where child i=1,2 (1= index child, 2=younger sibling) are siblings living in the same household k and ψ_k is the household fixed effect for the kth household. The advantage of this specification is that all unobserved characteristics that are common across siblings and that might bias the estimation of γ are purged from the estimation by including ψ_k. The application of household fixed effects relies on the assumption that γ is age invariant (e.g., Todd and Wolpin, 2003); this aspect is discussed when presenting the results.

5. Results

For all the main results presented in this section that test the association between a given FCS with PPVT, reading comprehension, math, and highest grade attained, in addition to standard p-values, we report adjusted p-values to adjust for multiple hypothesis testing based on Benjamini and Hochberg (1995). This approach controls for an expected proportion of falsely rejected hypotheses, and thus allows inference when conducting many tests (Anderson, 2008). We grouped all main outcomes (PPVT score, math score, reading comprehension score) for each RACER task, as they are analyzed individually. Specifically, we applied the multiple-test Simes (1986) procedure[vii]. We note below any discrepancies between the p-values and the q-values obtained as a result of the multiple hypothesis testing.

Table 4 presents the results for the main model (Equation 1) to assess the one-at-a-time relationships between each of the four FCS measured at age 12 and each cognitive-achievement-test score measured at age 15 (columns 1 to 3) as well as grade attainment by age 15 (column 4). The results are reported separately for Ethiopia (Panel A) and Peru (Panel B)[viii].

The results show that long-term memory and working memory are consistently and positively associated with the three measures of domain-specific cognitive achievements in both country samples. An increase of one standard deviation in long-term memory is associated with increases in cognitive achievement between 10% and 15% (12% and 14%) of a standard deviation in Ethiopia (Peru), whereas for working memory a one-standard-deviation increase is associated with increases in cognitive achievement between 8% and 14% (14% and 19%) of a standard deviation in Ethiopia (Peru).

INSERT TABLE 4 HERE

For inhibitory control, the signs of the point estimates are as expected but the coefficient estimates are not always statistically significant. A one-standard-deviation higher inhibitory control is associated significantly with increases in math scores in Ethiopia and Peru (by 13% and 12%, respectively); and, in PPVT in Ethiopia (by 8%). Finally, for implicit learning, one-standard-deviation higher scores are associated with higher PPVT and math scores in Ethiopia, by 15% and 10%, respectively. In the case of Peru, there are no significant associations observed.

For the highest grade attained by age 15, the results for Ethiopia show that the four FCSs are positively and significantly associated with grade attainment, with coefficient estimates that range between 0.26 and 0.35. In the case of Peru, only long-term memory and working memory are positively and significantly associated with the grade attained at age 15, with coefficient estimates of 0.05 and 0.17, respectively. Except for two coefficient estimates that were marginally significant, results in Table 4 are robust to multiple hypothesis testing.

In order to assess the predictive power of each FCS partialing-out the associations with the other FCSs, in Table A.1 (Appendix A) we include all the FCSs simultaneously, retaining the same adjustment variables included in Equation (1). There are four noteworthy results. First, virtually all the point estimates are smaller in magnitude, suggesting that the associations reported in Table 4, when only one of the four FCS components is included, partially capture the correlations with the other three FCS components. Second, for the standardized tests, all the coefficient estimates related to long-term memory and working memory remain statistically significant in both country samples. Third, inhibitory control retains significant predictive value only for math achievement in Ethiopia and Peru. In contrast, the significant association with PPVT is no longer observed for Ethiopia. Fourth, implicit learning no longer significantly predicts cognitive achievement in Ethiopia; and in Peru, counterintuitive results (associations with negative signs) are observed for math and reading comprehension, albeit with marginal significance. For grade attainment, all the coefficient estimates remain statistically significant for Ethiopia, whereas for Peru the coefficient estimates remain statistically significant for working memory and inhibitory control (marginally significant for the latter).

For PPVT and reading comprehension, the results could be driven by differences in the languages in which the test was administered. For instance, in the Peruvian sample, non-Spanish speakers perform worse in all RACER tasks compared to Spanish speakers (Behrman et al., 2022). To account for this while at the same taking into consideration that these tests were administered in the language spoken by the adolescents, in Table A.2 and Table A.3 (Appendix A) we report results separately by the main languages observed in each sample: Amharic, Tigrinya, and Oromo in Ethiopia; and Spanish in Peru. Overall, our conclusions remain unchanged.[ix]

It is conceivable that our main results are picking up a persistent relationship between FCSs and cognitive achievement over early and middle childhood, i.e., we might be capturing the persistent importance that each FCS had for cognitive achievement and grade attainment in younger ages. To establish whether this is the case, in Table 5 we report results of a value-added version of the main model (Equation 2). Thus, when the dependent variable is the highest grade attained observed at age 15, the estimation controls for highest grade attained at age 12. As a reference, the coefficient estimates linked to the lagged outcomes (ρ) are reported in Table A.4 (Appendix A). Comparing these results to the main results in Table 4, the new coefficient estimates are substantially smaller, although still statistically significant in most cases—9 out of 14 coefficients remain significant in Ethiopia, 8 out of 9 in Peru. This suggests that FCSs at age 12 play roles in explaining test scores and grade achievement at age 15 even when conditioning on those outcomes measured at age 12. The ρ coefficient for grade achievement is close to one, suggesting that FCSs explain changes in grades achieved between ages 12 and 15, whereas for test scores ρ is close to 0.5. The results highlight the importance of long-term memory and working memory to predict outcomes in all the observed domains of cognitive achievement, and inhibitory control to predict math results. For grade attainment, the relative declines of the coefficient estimates are much more noticeable; however, even for grade attainment long-term memory (in Ethiopia) and working memory (in both countries) remain significant predictors of changes in grade attainment. In Table A.5 (Appendix A), we include estimates for the value-added specification including all FCSs simultaneously.

INSERT TABLE 5 HERE

For the value-added specification, results are more sensitive to the adjustment for multiple hypothesis testing: in addition to some coefficient estimates that were only marginally significant in Ethiopia becoming insignificant (four coefficient estimates that link FCSs to PPVT and grade achievement), a few other coefficient estimates in Peru become insignificant (e.g., two coefficient estimates that link FCSs to PPVT). Those coefficient estimates linking FCSs to math and reading comprehension appear as the most robust in the two countries.

Up to this point, all outcomes that have been analyzed were observed in the last visit to the younger cohorts in 2016. We now incorporate outcomes observed during the phone survey in 2020 and 2021. Using this information, it is possible to identify whether YL participants had completed lower-secondary school in Ethiopia or were attending university in Peru when aged approximately 19 and 20 years old, respectively. Results must be interpreted considering that not all participants were contacted during this period. Notwithstanding this limitation, the relationship between FCSs at age 12 and having completed lower-secondary school (in Ethiopia) or university enrolment (in Peru) was explored through linear probability models including the same control variables used in the main model. The results for Ethiopia and Peru are presented in Table 6. In Ethiopia (column (1)), an improvement of one standard deviation in working memory predicts an increase of 3.4 percentage points in the probability of completing lower secondary school. In Peru (column (2)), a one-standard-deviation higher long-term memory is associated with an increase in the probability of ever being enrolled in university by 5.0 percentage points, whereas an analogous higher inhibitory control is associated with an increase in the probability of ever being enrolled in university by 3.9 percentage points. These results are consistent with the main results

(Table 4 and Table 5) that show that skills related to memory have the most relevant associations with educational achievement, while also highlighting the role of inhibitory control in Peru.

INSERT TABLE 6 HERE

6. Robustness checks

To understand potential attrition biases (e.g. respondents with outdated or missing contact information were not called), we made a comparison of average performance in each FCS of those YL participants contacted in the phone survey with those not contacted but that were observed in Round 5 (see Table A.6, Appendix A). We find that those not contacted during the phone survey have lower scores in all RACER tasks, compared to those that were contacted. For working memory and inhibitory control, the difference is statistically significant in both country samples. Assuming that those participants are also likely to have a lower educational attainment by age 19-20, these suggests our results in tables 7 and 8 are likely to represent lower bounds of the relationship between FCS and enrolment in higher education.

To deal with potential omitted household-level variable bias, Table 7 reports the results of re-estimating the main model including household fixed effects (Equation (3)), considering the pooled sample of index children and younger siblings—who are about three years younger, on average. Due to data availability, this model can only be estimated for PPVT and highest grade attained. To facilitate comparisons, columns (1) and (4) replicate the main findings from Table 4 (for PPVT and highest grade attained, respectively) for the sample of index children, whereas columns (2) and (5) present comparable results for the sample of younger siblings. For Ethiopia, positive associations between long-term memory, inhibitory control and implicit learning with PPVT remain significant in the sample of younger siblings, whereas the association with working memory becomes statistically insignificant. For Peru, the results show the same positive relationships between working memory and long-term memory with PPVT in the sample of younger siblings, and additionally we see an association between implicit learning and PPVT. In relation to grade attainment, for Ethiopia the predictive role of FCSs for grade attainment substantially declines for the younger siblings and only the one for long-term memory remains statistically significant. In contrast, for Peru the same relationships between working memory and long-term memory are observed for the younger siblings, and in this case a positive relationship between inhibitory control and grade attainment is also observed. Despite the differences observed in the point estimates obtained for the two groups of children, there does not appear to be a systematic difference (i.e., in a specific direction) between the coefficient estimates obtained for the two groups. An exception is grade attainment for Ethiopia, in which case the point estimates are substantially smaller for the younger siblings.

Results from the household fixed-effects specification are presented in columns (3) and (6)—for PPVT and highest grade attained, respectively. These results confirm that working memory and long-term memory predict cognitive achievement in vocabulary in Ethiopia and Peru, and that inhibitory control predicts vocabulary test scores in Ethiopia, whereas the results originally observed for implicit learning appear possibly to be spurious. For grade attainment, evidence is mixed. For Ethiopia, the coefficient estimates remain statistically significant for working memory and inhibitory control, but not for long-term memory. For Peru, the coefficient estimates linking working memory and long-term memory to grade attainment are smaller in magnitudes and become statistically insignificant. Results with household fixed effects are not sensitive to the adjustment for multiple hypothesis testing—only one coefficient estimate becomes marginally insignificant. Overall, these findings provide robust evidence of the associations between FCSs and test scores in both countries, and mixed evidence for the associations between FCSs and grade attainment.

INSERT TABLE 7 HERE

Previously we presented evidence from a value-added specification to gain insights on the contribution of FCSs at age 12 to predict changes in educational outcomes between ages 12 and 15. These results could also be afflicted by omitted variable bias due to unobserved household characteristics (Todd and Wolpin, 2003). To explore this, in Table 8 we apply household-fixed-effects to the value-added specification. As a reference, in columns (1) and (2) we report the value-added model separately for the index children and the younger siblings when using PPVT as an outcome, and in columns (4) and (5) the same is done for grade attainment.

The value-added model with household fixed effects shows meaningful associations between long-term memory and inhibitory control with PPVT in Ethiopia, whereas the association between working memory and PPVT becomes smaller and loses statistical significance (Table 8, column (3)). In the case of Peru, the associations between long-term memory and working memory with PPVT also become smaller and lose statistical significance (Table 8, column (3)), which might be due to precision loss as the point estimates still appear relevant. Finally, for grade attainment we find that the associations for Ethiopia remain statistically significant, whereas for Peru the coefficient estimates become insignificant—this pattern is similar to that observed in the model with only household fixed effects shown in Table 7.

INSERT TABLE 8 HERE

7. Discussion

In this study, we focus on the predictive power of foundational cognitive skills in late childhood at age 12 for educational outcomes at age 15; on the same results conditioning on those outcomes at age 12; as well as foundational cognitive skills predictive roles for completing lower-secondary school or entry into higher education. From a policy perspective, skill development and accessing post-secondary education are key areas of interest. Overall, we find that FCSs significantly predict later educational successes, which is consistent with the evidence from HICs. To our knowledge, our evidence using longitudinal data is unique for LMICs contexts. Additionally, our analysis focuses on FCSs of children in late childhood, for which ages prior evidence regarding the relationship between FCS and educational outcomes is scarce. We find that long-term memory and working memory are the only skills among those that we examine that significantly predict subsequent achievements across all domains: math, vocabulary, and reading comprehension. These FCSs also predict grade attainment in Ethiopia with results less robust for Peru. In contrast, implicit learning seems to play a minor role in predicting educational success in both countries. This is not surprising, since implicit learning is linked to very basic aspects of development during infancy (Amso & Davidow, 2012).

Our results also suggest significant associations in both countries between inhibitory control and math-test performance, but associations with vocabulary development and reading comprehension are less clear. The importance of inhibitory control for math achievement is consistent with prior evidence from HICs (Merkley et al., 2016; Clark et al., 2013). According to Merkley et al. (2016), this might be because numerical tasks require an ability to inhibit non-numerical stimuli. Despite its seemingly limited roles, inhibitory control is a significant predictor of grade attainment at age 15 in Ethiopia and university enrolment at age 20 in Peru. The associations between inhibitory control and grade attainment likely are mediated by its contribution to the development of math abilities, but inhibitory control also might play roles in non-cognitive dimensions that matter for school success, such as decisions around the time use of children (at school and at home) and risky behaviors (Blair et al., 2008). For example, by providing regulation, children might be more inclined to dedicate time to schoolwork, while this also impacts aspects like impulse and rage control.

Furthermore, simultaneously controlling for the four FCSs measured in RACER suggests that both long-term memory and working memory remain relevant to predict variation in test scores across all domains (math, vocabulary, and reading comprehension), and the associations of inhibitory control with math achievement persist. These results hold for both Ethiopia and Peru. The importance of multiple skills for school success is also reflected in the observed significant predictors of grade attainment. The two measures of executive functioning included in RACER—working memory and inhibitory control—appear relevant in both countries. The only meaningful difference observed between countries is that long-term memory does not predict grade attainment in Peru once the other FCSs are controlled for. Implicit learning has less consistent associations with educational outcomes.

From a theoretical point of view, an interesting question is why some cognitive skills are associated with educational performance or outcomes, and some are not. In the case of memory-related skills, long-term memory is associated with the ability to store new knowledge, while working memory is the ability to use that stored knowledge to complete tasks (Behrman et al., 2022). Those definitions are consistent with the requirements for successful continuous learning such as that children hopefully experience at school. So, it is not surprising that higher levels of long-term memory and working memory predict higher scores in learning outcomes and higher-grade attainment. Inhibitory control is partly related to the ability to focus on the task at hand by suppressing distracting stimuli while learning, which is another necessary tool to being able to perform well in learning activities. We can speculate that this is why inhibitory control has a significant association with some of the outcomes of interest. Finally, implicit learning is the ability to respond to regularities in the environment without necessarily being aware of them. The fact that we did not find this cognitive skill to be associated with educational outcomes in the two countries does not necessarily mean that it is not relevant in other outcomes or at earlier stages of development.

Another important aspect of our results is the evidence provided by the value-added specification about the role of FCSs in educational outcomes during adolescence conditional on those outcomes during late childhood. Like in the main model, in the value-added specification we observe that the memory-related FCSs are significantly associated with all three domain-specific test scores, PPVT, math and reading comprehension. Moreover, as observed for our main specification, inhibitory control is still associated with only the math test. As was mentioned, this evidence about the important roles of FCSs in changes in domain-specific skills between ages 12 and 15 is one of the main contributions of our paper. Commonly, educational policies focus on younger children in early childhood, but do not offer much help for older children that grew up with the challenges of LMICs settings. Therefore, the importance of our findings stems from what they say about the potential role of remediation policies aimed at older children. This aspect of our findings is also consistent with the potential malleability of CS, previously discussed. Nevertheless, it is important to make a qualification regarding this topic. According to the literature, ideally, cognitive malleability can be obtained through cognitive training, which can translate into the application of cognitive tasks (see Moreau, 2021). However, part of the evidence regarding the malleability of CS is mixed and based on the assessment of brief interventions designed to improve specific cognitive skills or intelligence in general. In relation to this, our results support the potential of improving a subgroup of these FCS skills, even well after early childhood. Nonetheless, this should be carefully interpreted as we are not performing any type of remediation or intervention, but only indicating the possibility of one being successful due to the positive associations we find between FCSs measured at age 12 and changes in the educational outcomes obtained between ages 12 and 15.

While we are unable to establish causal links without strong assumptions, the household fixed-effects strategy allows us to purge any bias derived from unobserved household characteristics that are constant over time, including (but not limited to) the role of differences in the quality of parenting. Although we cannot implement this specification for math and reading comprehension, as these tests were not administered to siblings, for vocabulary development our estimates confirm the importance of inhibitory control, working memory and long-term memory for predicting learning. When applying this strategy, it appears that cognitive skills are no longer relevant to predict grade attainment in Peru; however, this might be an artifact of limited within-household variation in grade attainment for this sample.

Our analysis has some limitations. First, since we observe FCSs during late childhood, at an age at which these measurements could already be affected by access to low-quality parenting and schools during the elementary-school and pre-school ages. However, we do not expect this to alter our conclusions because the value-added results show that FCSs still are associated with changes in educational outcomes between ages 12 and 15, and the household fixed-effects specification controls for school experiences that are common across siblings. Second, although our strategy allows us to deal with unobserved household heterogeneity, there might be unobserved child-level characteristics that could bias our results (e.g., educational influences coming from sources other than the exposures shared by siblings in schools or families). Therefore, we are able to make causal claims about the relationships observed only conditional on unobserved child characteristics that are not controlled by the baseline scores and the lagged scores in the value-added estimates not being very important. Third, the long-term analysis using data from the phone survey may be afflicted by attrition bias, as we were unable to reach participants for whom we did not have up-to-date contact data, and those affected by the civil war in the Tigray region in Ethiopia. This might lead to a downward bias estimation as the selection would tend to exclude children from poorer backgrounds who arguably would have had larger effects from higher FCSs.

From a policy perspective, our results justify the importance of investing in cognitive skills that are relevant for educational success. In particular, the results presented show robust evidence of the importance of working memory and inhibitory control—two key areas of executive functioning, as well as long-term memory for predicting learning outcomes in LMICs contexts. Furthermore, our results show that FCSs are relevant to predict higher-education enrollment many years later. This poses a challenge for school systems: children need to develop FCSs that allow them to perform well at school tasks (e.g., school assignments) and continue learning; at the same time, school tasks can enhance FCSs. This virtuous circle can become a vicious circle for children from vulnerable backgrounds that start schools without a minimum set of FCSs and fall behind early in life. As shown by our results in LMICs and those from Jerrim and Vignole (2015) in HICs, the impact of low CSs development can go as far as to limit post-secondary education achievement.

Initiatives related to designing an educational plan more focused on the development of CSs, rather than just knowledge impartment, should not only be implemented for the first years of education, but also be extended for older children, working as a remediation policy. Banerjee et al. (2007) provides good examples of the potential of cost-effective remediation policies to increase children´s performance. Both programs evaluated by the authors are found to have gains in the average scores in literacy and numeracy skills of Indian children in elementary grade levels. These results are consistent with what Muralidharan et al. (2019) find in their evaluation of another program implemented in India, with both studies highlighting how children low in the score distribution benefit most. The differential effects of interventions to improve cognitive skills at different stages of human development in LMICs remains an interesting topic for further research. At the end, both early and later interventions are potentially important investments that can help disadvantaged children achieve their potentials.

References

Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. Journal of the American Statistical Association, 103(484), 1481–1495. https://doi.org/10.1198/016214508000000841

Amso, D., & Davidow, J. (2012). The development of implicit learning from infancy to adulthood: item frequencies, relations, and cognitive flexibility. *Developmental Psychobiology*, *54*(6), 664–673. https://doi.org/10.1002/dev.20587

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *The Quarterly Journal of Economics*, *131*(3), 1415–1454. https://www.jstor.org/stable/26372667

Banerjee, A. V., Cole, S., Duflo, E., & Linden, L. (2007). Remedying Education: Evidence from Two Randomized Experiments in India. *The Quarterly Journal of Economics, 122*(3), 1235-1264.

Benes, F. M. (2006). The development of the prefrontal cortex: The maturation of neurotransmitter systems and their interactions. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental Psychopathology: Developmental Neuroscience* (pp. 216–258). John Wiley & Sons Inc.

Behrman, J.R., Briones, K., Cueto, S., Favara, M., Freund, R, Hittmeyer, A., Lopez, J., Pazos Navarro, N., Sánchez, A., Scott, D., Sheridan, M., & Woldehanna, T. (2022). *Measuring foundational cognitive skills in Young Lives using RACER*. (Young Lives Technical Note 54). Young Lives.

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child development, 81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Blair C. (2002). School readiness. Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *The American Psychologist, 57*(2), 111–127. https://doi.org/10.1037//0003-066x.57.2.111

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*(2), 647–663. https://doi.org/10.1111/j.1467-8624.2007.01019.x

Blair, C., & Diamond, A. (2008). Biological processes in prevention and intervention: the promotion of self-regulation as a means of preventing school failure. *Development and Psychopathology*, *20*(3), 899–911. https://doi.org/10.1017/S0954579408000436

Blair, C., Knipe, H., & Gamson, D. (2008). Is there a role for executive functions in the development of mathematics ability? *Mind, Brain, and Education, 2*(2), 80–89. https://doi.org/10.1111/j.1751-228X.2008.00036.x

Briones, K. 2017. How many rooms are there in your house? Constructing the young lives wealth index. Young Lives Technical Note, 43, 1–18.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33*(3), 205–228. https://doi.org/10.1080/87565640801982312

Bull, R., Espy, K. A., Wiebe, S. A., Sheffield, T. D., & Nelson, J. M. (2011). Using confirmatory factor analysis to understand executive control in preschool children: sources of variation in emergent mathematic achievement. *Developmental Science, 14*(4), 679–692. https://doi.org/10.1111/j.1467-7687.2010.01012.x

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's Reading Comprehension Ability: Concurrent Prediction by Working Memory, Verbal Ability, and Component Skills. *Journal of Educational Psychology, 96*(1), 31–42. https://doi.org/10.1037/0022-0663.96.1.31

Cantor, P., Osher, D., Berg, J., Steyer, L., & Rose, T. (2019). Malleability, plasticity, and individuality: How children learn and develop in context. *Applied Developmental Science, 23*(4), 307–337. https://doi.org/10.1080/10888691.2017.1398649

Carlson S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology, 28*(2), 595–616. https://doi.org/10.1207/s15326942dn2802_3

Carneiro, P., & Heckman, J. J. (2002). The Evidence on Credit Constraints in Post-secondary Schooling. *Economic Journal, 112*(482), 705-734.

Castro, J. F., Yamada, G., & Arias, O. (2016). Higher education decisions in Peru: on the role of financial constraints, skills, and family background. *Higher Education, 72*(4), 457–486. http://www.jstor.org/stable/26447555

Chen, A., Panter-Brick, C., Hadfield, K., Dajani, R., Hamoudi, A., & Sheridan, M. (2019). Minds under siege: Cognitive signatures of poverty and trauma in refugee and nonrefugee adolescents. *Child Development*, 90(6), 1856–1865.

Christopher, M. E., Miyake, A., Keenan, J. M., Pennington, B., DeFries, J. C., Wadsworth, S. J., Willcutt, E., & Olson, R. K. (2012). Predicting word reading and comprehension with executive function and speed measures across development: a latent variable analysis. *Journal of Experimental Psychology. General*, *141*(3), 470–488. https://doi.org/10.1037/a0027375

Clark, C. A. C., Sheffield, T. D., Wiebe, S. A., and Espy, K. A. (2013). Longitudinal associations between executive control and developing mathematical competence in preschool boys and girls. *Child Development, 84*, 662–677. doi: 10.1111/j.1467-8624.2012.01854.x

Cueto, S., León, J., Guerrero, G., & Muñoz, I. (2009). Psychometric characteristics of cognitive development and achievement instruments in Round 2 of Young Lives (Young Lives Technical Note 15). Young Lives.

Cueto, S., & Leon, J. (2012). Psychometric characteristics of cognitive development and achievement instruments in Round 3 of Young Lives (Young Lives Technical Note 25). Young Lives

Cunha, F., Heckman, J., Lochner, L., & Masterov, D. (2006). Interpreting the evidence on life cycle skill formation. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the economics of education, Chapter 12* (pp. 697–812). Amsterdam: North-Holland.

Dilworth-Bart, J. E. (2012). Does executive function mediate SES and home quality associations with academic readiness? *Early Childhood Research Quarterly, 27*(3), 416–425. https://doi.org/10.1016/j.ecresq.2012.02.002

Escobal, J., & Flores, E. (2008). An assessment of the Young Lives sampling approach in Peru. Young Lives Technical Note 3, Oxford: Young Lives. Available from: www.younglives.org.uk/publications

Evans, D. K., & Popova, A. (2016). What Really Works to Improve Learning in Developing Countries? An Analysis of Divergent Findings in Systematic Reviews. *The World Bank Research Observer*, *31*(2), 242–270. https://doi.org/10.1093/wbro/lkw004

Ford, C. B., Kim, H. Y., Brown, L., Aber, J. L., & Sheridan, M. A. (2019). A cognitive assessment tool designed for data collection in the field in low- and middle-income countries. Research in Comparative and International Education, 14(1), 141–157.

Freund, R., Favara, M., Porter, C., & Behrman, J. (2023). Social Protection and foundational cognitive skills during adolescence: Evidence from a large public works program. *The World Bank Economic Review*.

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology. General*, *137*(2), 201–225. https://doi.org/10.1037/0096-3445.137.2.201

Friedman, N. P., Miyake, A., Robinson, J. L., & Hewitt, J. K. (2011). Developmental trajectories in toddlers' self-restraint predict individual differences in executive functions 14 years later: a behavioral genetic analysis. *Developmental Psychology, 47*(5), 1410–1430. https://doi.org/10.1037/a0023750

Friedman, N. P., Miyake, A., Altamirano, L. J., Corley, R. P., Young, S. E., Rhea, S. A., & Hewitt, J. K. (2016). Stability and change in executive function abilities from late adolescence to early adulthood: A longitudinal twin study. *Developmental Psychology, 52*(2), 326–340. https://doi.org/10.1037/dev0000075

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychological Bulletin, 134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Geary, D. C. (2013). Early foundations for mathematics learning and their relations to learning disabilities. *Current Directions in Psychological Science, 22*(1), 23–27. https://doi.org/10.1177/0963721412469398

Glewwe, P., Hanushek, E., Humpage, S., & Ravina, R. (2011). *School Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010*. (National Bureau of Economic Research, Working Papers 17554). https://www.nber.org/papers/w17554

Griffin, E. A., & Morrison, F. J. (1997). The unique contribution of home literacy environment to differences in early literacy skills. *Early Child Development and Care*, *127-128,* 233–243. https://doi.org/10.1080/0300443971270119

Hamoudi, A., & Sheridan, M.A. (2015). Unpacking the Black Box of Cognitive Ability A novel tool for assessment in a population based survey.
Heckman, J. J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science, 312*(5782), 1900–1902. doi:10.1126/science.1128898
Howse, R. B., Lange, G., Farran, D. C., & Boyles, C. D. (2003). Motivation and self-regulation as predictors of achievement in economically disadvantaged young children. *Journal of Experimental Education, 71*(2), 151–174.
Hanushek, E. A. (2009). The Economic Value of Education and Cognitive Skills. In G. Sykes, B. Schneider, & D. N. Plank (Eds.), *Handbook of Education Policy Research* (pp. 39–56).
Hanushek, E. A., & Woessmann, L. (2008). The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature, 46*(3), 607–668.
Huizinga, M., Dolan, C. V., & van der Molen, M. W. (2006). Age-related change in executive function: developmental trends and a latent variable analysis. *Neuropsychologia, 44*(11), 2017–2036. https://doi.org/10.1016/j.neuropsychologia.2006.01.010
Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences of the United States of America, 108*(25), 10081–10086. https://doi.org/10.1073/pnas.1103228108
Jerrim, J., & Vignoles, A. (2016). The link between East Asian ´mastery´teaching methods and English children´s mathematics skills. *Economics of Education Review, 50*(4), 29–44. https://doi.org/10.1016/j.econedurev.2015.11.003
Klingberg T. (2010). Training and plasticity of working memory. Trends in Cognitive Sciences, 14(7), 317–324. https://doi.org/10.1016/j.tics.2010.05.002
Knoll, L. J., Fuhrmann, D., Sakhardande, A. L., Stamp, F., Speekenbrink, M., & Blakemore, S.-J. (2016). A Window of Opportunity for Cognitive Training in Adolescence. *Psychological Science, 27*(12), 1620–1631. https://doi.org/10.1177/0956797616671327
Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: a cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology, 108*(3), 677–692. https://doi.org/10.1016/j.jecp.2010.11.001
Lee, K., Bull, R., & Ho, R. M. (2013). Developmental changes in executive functioning. *Child Development, 84*(6), 1933–1953. https://doi.org/10.1111/cdev.12096
McAuley, T., & White, D. A. (2011). A latent variables examination of processing speed, response inhibition, and working memory during typical development. *Journal of Experimental Child Psychology, 108*(3), 453–468. https://doi.org/10.1016/j.jecp.2010.08.009
McClelland, M. M., Morrison, F. J., & Holmes, D. L. (2000). Children at risk for early academic problems: The role of learning related social skills. *Early Childhood Research Quarterly*, 15, 307-329. http://dx.doi.org/10.1016/S0885-2006(00)00069-7
McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., & Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Developmental Psychology, 43*(4), 947–959. https://doi.org/10.1037/0012-1649.43.4.947
McEwan, P. J. (2015). Improving Learning in Primary Schools of Developing Countries: A Meta-Analysis of Randomized Experiments. *Review of Educational Research, 85*(3), 353–394. https://doi.org/10.3102/0034654314553127
Merkley, R., Thompson, J., & Scerif, G. (2016). Of Huge Mice and Tiny Elephants: Exploring the Relationship Between Inhibitory Processes and Preschool Math Skills. *Frontiers in Psychology, 6*, 1903. https://doi.org/10.3389/fpsyg.2015.01903
Micalizzi, L., Brick, L. A., Flom, M., Ganiban, J. M., & Saudino, K. J. (2019). Effects of socioeconomic status and executive function on school readiness across levels of household chaos. *Early Childhood Research Quarterly, 47*, 331–340. https://doi.org/10.1016/j.ecresq.2019.01.007
Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool

children. *Journal of Cognition and Development*, *13*(3), 395–423. https://doi.org/10.1080/15248372.2011.585478

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive Psychology, 41*(1), 49–100. https://doi.org/10.1006/cogp.1999.07347

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Monette, S., Bigras, M., & Lafrenière, M. A. (2015). Structure of executive functions in typically developing kindergarteners. *Journal of Experimental Child Psychology*, *140*, 120–139. https://doi.org/10.1016/j.jecp.2015.07.005

Moreau, D. (2022). How malleable are cognitive abilities? A critical perspective on popular brief interventions. *American Psychologist, 77*(3), 409–423. https://doi.org/10.1037/amp0000872

Muralidharan, K., Singh, A., & Ganimian, A. J. (2019). Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India. *American Economic Review, 109*(4), 1426–1460. doi:10.1257/aer.20171112

Outes-Leon, I., & Sánchez, A. (2016). An assessment of the Young Lives sampling approach in Ethiopia (Updated Publisher's version, Young Lives Technical Note). Young Lives.

Reynolds, A. J., Temple, J. A., & Ou, S. R. (2010). Preschool Education, Educational Attainment, and Crime Prevention: Contributions of Cognitive and Non-Cognitive Skills. *Children and Youth Services Review, 32*(8), 1054–1063. https://doi.org/10.1016/j.childyouth.2009.10.019

Riggs, N. R., Black, D. S., & Ritt-Olson, A. (2015). Associations between dispositional mindfulness and executive function in early adolescence. *Journal of Child and Family Studies, 24*(9), 2745–2751. https://doi.org/10.1007/s10826-014-0077-3

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The Power of Personality: The Comparative Validity of Personality Traits, Socioeconomic Status, and Cognitive Ability for Predicting Important Life Outcomes. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *2*(4), 313–345. https://doi.org/10.1111/j.1745-6916.2007.00047.x

Sánchez, A., Favara, M., Sheridan, M., & Behrman, J. (2024). "Does early nutrition predict cognitive skills during later childhood? Evidence from two developing countries," *World Development.*

Scott, D., Lopez, J., Sánchez, A., & Behrman, J. (2022). The impact of the JUNTOS conditional cash transfer programme on foundational cognitive skills: Does age of enrolment matter? (Penn Institute for Economics Research (PIER) No 22-019)

Simes, R. J. 1986. An improved Bonferroni procedure for multiple tests of significance. *Biometrika*, *73*, 751–754.

Schmitt, S. A., Geldhof, G. J., Purpura, D. J., Duncan, R., & McClelland, M. M. (2017). Examining the relations between executive function, math, and literacy during the transition to kindergarten: A multi-analytic approach. *Journal of Educational Psychology*, *109*(8), 1120–1140. https://doi.org/10.1037/edu0000193

Stipek, D. J., & Ryan, R. H. (1997). Economically disadvantaged preschoolers: ready to learn but further to go. *Developmental Psychology, 33*(4), 711–723. https://doi.org/10.1037//0012-1649.33.4.711

Todd, P. & Wolpin, K. (2003). On the specification and estimation of the production function for cognitive of the production function for cognitive. *The Economic Journal, 113,* 3-33.

United Nations (2015). *Transforming Our World: The 2030 Agenda for Sustainable Development*. New York: UN.

Usai, M. C., Viterbori, P., Traverso, L., & De Franchis, V. (2014). Latent structure of executive function in five- and six-year-old children: A longitudinal study. *European Journal of*

Developmental Psychology, *11*(4), 447–462. https://doi.org/10.1080/17405629.2013.840578

Willoughby, M. T., Piper, B., Kwayumba, D., & McCune, M. (2019). Measuring executive function skills in young children in Kenya. Child Neuropsychology, 25(4), 425–444.

Wray, C., Kowalski, A., Mpondo, F., Ochaeta, L., Belleza, D., DiGirolamo, A., Waford, R., Richter, L., Lee, N., Scerif, G., Stein, A. D., & Stein, A. (2020). Executive functions form a single construct and are associated with schooling: Evidence from three low-and middle- income countries. PLoS ONE, 15(11), E0242936.

Yuan, H., Ocansey, M., Adu-Afarwuah, S., Sheridan, M., Hamoudi, A., Okronipa, H., Kumordzie, S. M., Oaks, B. M., & Prado, E. (2022). Evaluation of a tablet-based assessment tool for measuring cognition among children 4-6 years of age in Ghana. *Brain and Behavior*, *12*(10), e2749. https://doi.org/10.1002/brb3.2749

Zelazo, P. D., Lourenco, S. F., Frank, M. C., Elison, J. T., Heaton, R. K., Wellman, H. M., Slotkin, J., Kharitonova, M., & Reznick, J. S. (2021). Measurement of Cognition for the National Children's Study. *Frontiers in Pediatrics*, *9*, 603126. https://doi.org/10.3389/fped.2021.603126

Tables

Table 1: Summary statistics

	Ethiopia	Peru
Panel A		
Child lives in urban area (in %), round 5 (2016)	36.5%	74.7%
Child is male (in %)	52.9%	50.4%
Maternal schooling (in %)		
None	38.0%	n.a.
Any grade of elementary schooling	49.8%	n.a.
More than elementary schooling	12.2%	n.a.
None	n.a.	8.4%
Any grade of primary schooling	n.a.	36.4%
Any grade of secondary schooling	n.a.	37.1%
More than secondary schooling	n.a.	18.1%
Region of origin / ethnicity		
Mother´s native tongue is Amharic	34.6%	n.a.
Mother´s native tongue is Tigrinya	22.1%	n.a.
Mother´s native tongue is Oromifa	18.7%	n.a.
Mother´s native tongue is Spanish	n.a.	70.3%
Child´s native tongue is Amharic	42.4%	n.a.
Child´s native tongue is Tigrinya	21.8%	n.a.
Child´s native tongue is Oromifa	16.5%	n.a.
Child´s native tongue is Spanish	n.a.	85.6%
Panel B		
Age in round 5 (2016) (average)	15.1	14.9
Highest grade attained (average)	6.8	8.9
Grade attained in round 5 (2016)		
9th grade - Secondary schooling	18.4%	n.a.
8th grade - Elementary schooling	21.6%	n.a.
7th grade - Elementary schooling	13.5%	n.a.
6th grade and below	39.5%	n.a.
10th grade - Secondary schooling	n.a.	26.2%
9th grade - Secondary schooling	n.a.	48.7%
8th grade - Secondary schooling,	n.a.	14.7%
7th grade and below	n.a.	9.8%
Long-term educational outcomes (2020 Ethiopia, 2021 Peru)		
Finished secondary education	n.a.	90.0%
Finished lower secondary education	45.5%	n.a.
Finished upper secondary education	11.7%	n.a.
Ever enrolled in higher education	3.1%	58.8%
Ever enrolled in university	2.5%	31.8%
Panel C – response rates on cognitive tests		
Took the PPVT, round 5 (2016)	86.2%	95.7%
Took the math test, round 5 (2016)	91.3%	96.9%
Took the reading comprehension test, round 5 (2016)	89.9%	95.0%
Took the RACER, round 4 (2013)	99.9%	99.0%
Panel D – language of administration (PPVT, math, reading comprehension)		
Amharic	48.1%	n.a.
Tigrinya	21.5%	n.a.
Oromifa	19.4%	n.a.
Spanish	n.a.	98.8%
Quechua	n.a.	0.1%
Spanish and Quechua	n.a.	0.2%

Note: statistics correspond to the sample of index children that are part of the analytical sample.

Table 2: Description of RACER tasks

RACER Task #	Cognitive Task	Cognitive Ability	Definition
Task 1	Paired Associate Learning Task (Part 1) "Memory Game 1"	Long-term Memory / Declarative memory	**Long-term memory / declarative memory**: the ability to encode and retain new knowledge.
Task 2	Simon Task "Sides Game"	Inhibitory Control	**Inhibition**: the ability to stop oneself from exhibiting behaviours one does not want to exhibit and is related to one's ability to focus on a single task and supress distractors.
Task 3	Spatial Delayed-Match-to-Sample Task "Finding the Dots"	Working Memory	**Working memory**: the ability to hold in mind and manipulate stimuli that are no longer present in the environment.
Task 4	Adapted Serial Reaction Time Task "Catching Chickens/ Chasing dots"	Implicit Learning	**Implicit learning**: the ability of the motor system to recognise and respond to regularities in the environment even when individuals are not aware of these regularities.
Task 5	Paired Associate Learning Task (Part 2) "Memory Game 2"	Long-term Memory / Declarative memory	**Long-term memory / Declarative memory:** the ability to encode and retain new knowledge.

Note: this table is taken from Behrman et al. (2022). The table is used with authorization from the authors.

Table 3: Summary statistics

	Inhibitory control		Working memory		Long-term memory		Implicit learning	
	Ethiopia	Peru	Ethiopia	Peru	Ethiopia	Peru	Ethiopia	Peru
Full sample	-0.130	0.127	-0.148	0.158	-0.134	0.126	-0.095	0.092
By area of location								
Urban	0.039	0.281	-0.017	0.296	0.158	0.213	0.093	0.167
Rural	-0.217	-0.211	-0.217	-0.168	-0.291	-0.067	-0.196	-0.072
By tertiles of household wealth								
Bottom	-0.190	-0.178	-0.291	-0.139	-0.245	-0.079	-0.200	-0.165
Middle	-0.242	0.095	-0.156	0.154	-0.327	0.132	-0.173	0.092
Top	0.054	0.466	0.007	0.442	0.181	0.323	0.080	0.351

Note: These results are reported for the balanced sample of participants. The results have been standardized to have mean zero and variance one by age in years within the pooled sample of Ethiopia and Peru.

Table 4:
Associations, one at a time, between each of the four foundational cognitive skills at age 12 and three test scores and highest grade attained at age 15 (Ordinary Least Squares)

Dependent variables:		**PPVT**	**Math**	**Reading comprehension**	**Highest grade attained**
		(1)	**(2)**	**(3)**	**(4)**
Panel A: Ethiopia					
Independent variables					
Long-term memory	Coef.	0.146***	0.095***	0.125***	0.287***
	s.e.	(0.038)	(0.026)	(0.024)	(0.057)
	q-value	0.002	0.004	0.000	0.000
Inhibitory control	Coef.	0.082*	0.128***	0.026	0.319***
	s.e.	(0.045)	(0.037)	(0.033)	(0.080)
	q-value	0.134	0.005	0.531	0.002
Working memory	Coef.	0.141***	0.111***	0.080***	0.260***
	s.e.	(0.023)	(0.025)	(0.024)	(0.062)
	q-value	0.000	0.001	0.009	0.001
Implicit learning	Coef.	0.153**	0.098*	0.092	0.351**
	s.e.	(0.071)	(0.049)	(0.054)	(0.136)
	q-value	0.083	0.102	0.162	0.038
Obs.		1469	1564	1541	1649
Adjusted R2		0.464	0.324	0.398	0.458
Adjusted R2		0.462	0.329	0.393	0.472
Adjusted R2		0.486	0.333	0.401	0.472
Adjusted R2		0.454	0.320	0.389	0.451
Panel B: Peru					
Independent variables					
Long-term memory	Coef.	0.144***	0.124***	0.141***	0.053**
	s.e.	(0.016)	(0.021)	(0.029)	(0.021)
	q-value	0.000	0.000	0.000	0.032
Inhibitory control	Coef.	0.064	0.124***	0.035	0.077
	s.e.	(0.042)	(0.042)	(0.042)	(0.046)
	q-value	0.196	0.017	0.456	0.153
Working memory	Coef.	0.193***	0.182***	0.136***	0.167***
	s.e.	(0.021)	(0.024)	(0.027)	(0.035)
	q-value	0.000	0.000	0.000	0.000
Implicit learning	Coef.	0.048	0.026	0.027	0.050
	s.e.	(0.030)	(0.025)	(0.021)	(0.030)
	q-value	0.169	0.372	0.277	0.153
Obs.		1697	1718	1686	1718
Adjusted R2		0.330	0.203	0.203	0.260
Adjusted R2		0.341	0.216	0.205	0.293
Adjusted R2		0.359	0.244	0.229	0.299
Adjusted R2		0.314	0.194	0.185	0.260

Note: Each coefficient in each cell comes from a different estimated model. Each model controls for the baseline performance in the associated task, child´s sex, age in months, ethnicity, and number of years of pre-schooling; household wealth dummies (1 if in the middle tertile, 0 otherwise; 1 if in the top tertile, 0 otherwise), maternal level of schooling, area of residence, time of RACER administration (1 morning, 0 otherwise), date of RACER administration (1 if on a week day, 0 otherwise), and community-of-birth fixed effects. ***p<0.01, **p<0.05, *p<0.10. q-values correct for multiple hypothesis testing.

Table 5:
Associations, one at a time, between foundational cognitive skills at age 12 and test scores and highest grade attained at age 15 (Value-added specifications)

Dependent variables:		**PPVT**	**Math**	**Reading comprehension**	**Highest grade attained**
		(1)	**(2)**	**(3)**	**(4)**
Panel A: Ethiopia					
Independent variables					
Long-term memory	Coef.	0.048*	0.023	0.071***	0.040*
	s.e.	(0.027)	(0.028)	(0.019)	(0.021)
	q-value	0.169	0.552	0.004	0.140
Inhibitory control	Coef.	0.029	0.076***	-0.030	0.067
	s.e.	(0.036)	(0.026)	(0.027)	(0.052)
	q-value	0.564	0.021	0.403	0.329
Working memory	Coef.	0.066***	0.054**	0.053**	0.040*
	s.e.	(0.020)	(0.021)	(0.025)	(0.023)
	q-value	0.010	0.045	0.093	0.176
Implicit learning	Coef.	0.125*	0.013	0.051	0.059
	s.e.	(0.061)	(0.035)	(0.041)	(0.047)
	q-value	0.111	0.808	0.347	0.346
Obs.		1439	1414	1368	1649
Adjusted R2		0.606	0.478	0.525	0.843
Adjusted R2		0.606	0.480	0.521	0.843
Adjusted R2		0.616	0.481	0.526	0.844
Adjusted R2		0.607	0.478	0.521	0.843
Panel B: Peru					
Independent variables					
Long-term memory	Coef.	0.072***	0.050***	0.073***	-0.002
	s.e.	(0.018)	(0.016)	(0.021)	(0.015)
	q-value	0.933	0.002	0.012	0.008
Inhibitory control	Coef.	0.031	0.060**	-0.018	0.010
	s.e.	(0.028)	(0.027)	(0.034)	(0.025)
	q-value	0.768	0.384	0.062	0.679
Working memory	Coef.	0.093***	0.057**	0.058**	0.024*
	s.e.	(0.014)	(0.022)	(0.024)	(0.013)
	q-value	0.144	0.000	0.034	0.045
Implicit learning	Coef.	0.013	-0.026	-0.004	-0.005
	s.e.	(0.024)	(0.018)	(0.020)	(0.012)
	q-value	0.747	0.692	0.261	0.899
Obs.		1692	1710	1680	1718
Adjusted R2		0.549	0.482	0.400	0.793
Adjusted R2		0.550	0.481	0.397	0.793
Adjusted R2		0.554	0.486	0.408	0.793
Adjusted R2		0.545	0.482	0.395	0.793

Note: Each coefficient in each cell comes from a different estimated model. Each model controls for the baseline performance in all (four) tasks, lagged test score, child´s sex, age in months, ethnicity, and number of years of pre-schooling; household wealth dummies (1 if in the middle tertile, 0 otherwise; 1 if in the top tertile, 0 otherwise), maternal level of schooling, area of residence, time of RACER administration (1 morning, 0 otherwise), date of RACER administration (1 if on a week day, 0 otherwise), and community-of-birth fixed effects. ***p<0.01, **p<0.05, *p<0.10. q-values correct for multiple hypothesis testing.

Table 6:
Association one at a time between foundational cognitive skills at age 12 and educational attainment at age 19-20 (Ordinary Least Squares)

		(1)	(2)
		Ethiopia	Peru
Dependent variables:		***Finished lower secondary education at age 19 (1 if yes; 0 if no)***	***Enrolled in university at age 20 (1 if yes; 0 if no)***
Independent variables:			
Long-term memory	Coef.	0.0149	0.0502***
	s.e.	(1.66)	(4.49)
Inhibitory control	Coef.	0.0317	0.0390*
	s.e.	(1.57)	(2.36)
Working memory	Coef.	0.0335*	0.0274
	s.e.	(2.73)	(2.03)
Implicit learning	Coef.	0.0169	0.0228
	s.e.	(1.15)	(0.88)
Obs.		1543	1506

Note: In column (1) (Ethiopia), having finished lower secondary takes the value of 1 if the YL participant completed at least Grade 10. In column (2) (Peru), enrolment in university takes the value of 1 if the YL participant is currently enrolled or has been enrolled in a university institution, 0 otherwise. In columns (1) and (2), each coefficient in each cell comes from a different estimated model. Each model controls for the baseline performance in the associated task, child´s sex, age in months, ethnicity, and number of years of pre-schooling; household wealth dummies (1 if in the middle tertile, 0 otherwise; 1 if in the top tertile, 0 otherwise), maternal schooling level, area of residence, time of RACER administration (1 morning, 0 otherwise), date of RACER administration (1 if on a weekday, 0 otherwise), and community-of-birth fixed effects. ***p<0.01, **p<0.05, *p<0.10.

Table 7:
Household fixed-effects specification – Foundational cognitive skills at age 12 and educational outcomes at age 15

		Dep. Variable: PPVT			Dep. Variable: Highest grade attained		
		Index children (OLS)	**Younger sibling (OLS)**	**HFE**	**Index children (OLS)**	**Younger sibling (OLS)**	**HFE**
		(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Panel A: Ethiopia							
Independent variables:							
Long-term memory	Coef.	0.146***	0.123***	0.062***	0.287***	0.150**	0.067
	s.e.	(0.038)	(0.031)	(0.019)	(0.057)	(0.056)	(0.060)
	q-value	0.002	0.002	0.010	0.000	0.031	0.390
Inhibitory control	Coef.	0.082*	0.093*	0.116**	0.319***	0.043	0.433***
	s.e.	(0.045)	(0.046)	(0.042)	(0.080)	(0.100)	(0.136)
	q-value	0.134	0.096	0.024	0.002	0.741	0.011
Working memory	Coef.	0.141***	0.069	0.072**	0.260***	0.061	0.219***
	s.e.	(0.023)	(0.042)	(0.034)	(0.062)	(0.060)	(0.063)
	q-value	0.000	0,172	0.070	0.001	0.397	0.005
Implicit learning	Coef.	0.153**	0.097**	0.038	0.351**	0.006	0.164
	s.e.	(0.071)	(0.036)	(0.040)	(0.136)	(0.072)	(0.095)
	q-value	0.083	0.031	0.461	0.038	0.931	0.203
Obs.		1469	705	2196	1649	707	2481
Adjusted R2		0.464	0.428	0.077	0.458	0.563	0.720
Adjusted R2		0.462	0.439	0.094	0.472	0.566	0.733
Adjusted R2		0.486	0.431	0.100	0.472	0.570	0.730
Adjusted R2		0.454	0.418	0.071	0.451	0.557	0.721
Panel B: Peru							
Independent variables:							
Long-term memory	Coef.	0.144***	0.118***	0.069**	0.053**	0.065*	0.020
	s.e.	(0.016)	(0.035)	(0.032)	(0.021)	(0.036)	(0.034)
	q-value	0.000	0.011	0.104	0.032	0.139	0.867
Inhibitory control	Coef.	0.064	0.023	0.006	0.077	0.164**	0.065
	s.e.	(0.042)	(0.048)	(0.041)	(0.046)	(0.060)	(0.063)
	q-value	0.196	0.739	0.924	0.153	0.030	0.499
Working memory	Coef.	0.193***	0.134***	0.106***	0.167***	0.162***	0.067
	s.e.	(0.021)	(0.034)	(0.030)	(0.035)	(0.035)	(0.050)
	q-value	0.000	0.003	0.011	0.000	0.001	0.344
Implicit learning	Coef.	0.048	0.106***	0.001	0.050	-0.001	-0.026
	s.e.	(0.030)	(0.036)	(0.035)	(0.030)	(0.073)	(0.065)
	q-value	0.169	0.021	0.975	0.153	0.992	0.857
Obs.		1697	676	2381	1718	670	2398
Adjusted R2		0.330	0.404	0.027	0.260	0.653	0.888
Adjusted R2		0.341	0.407	0.042	0.293	0.659	0.889
Adjusted R2		0.359	0.435	0.039	0.299	0.660	0.889
Adjusted R2		0.314	0.397	0.022	0.260	0.652	0.889

Note: Each coefficient in each cell comes from a different estimated model. Results from columns (1), (2), (4) and (5) are based on OLS and follow Equation (1). Results from columns (1) and (4) —previously reported in Table 4—use data from the index children, and (2) and (5) from the younger siblings. Results from columns (3) and (6) correspond to the household fixed effects model represented in Equation (3). ***p<0.01, **p<0.05, *p<0.10. q-values correct for multiple hypothesis testing.

Table 8:
Value-added specification without and with household fixed-effects – Foundational cognitive skills at age 12 and educational outcomes at age 15

		Dep. Variable: PPVT			Dep. Variable: Highest grade attained		
		Index children (VA)	**Younger sibling (VA)**	**VA-HFE**	**Index children (VA)**	**Younger sibling (VA)**	**VA-HFE**
		(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Panel A: Ethiopia							
Independent variables:							
Long-term memory	Coef.	0.048*	0.062**	0.058***	0.040*	0.150**	-0.007
	s.e.	(0.027)	(0.022)	(0.018)	(0.021)	(0.056)	(0.047)
	q-value	0.169	0.035	0.011	0.140	0.751	0.905
Inhibitory control	Coef.	0.029	0.085*	0.101*	0.067	0.043	0.277**
	s.e.	(0.036)	(0.045)	(0.037)	(0.052)	(0.100)	(0.099)
	q-value	0.564	0.142	0.028	0.329	0.662	0.025
Working memory	Coef.	0.066***	0.004	0.047	0.040*	0.061	0.115*
	s.e.	(0.020)	(0.038)	(0.032)	(0.023)	(0.060)	(0.055)
	q-value	0.010	0.936	0.211	0.176	0.679	0.095
Implicit learning	Coef.	0.125*	0.067	0.026	0.059	0.006	0.047
	s.e.	(0.061)	(0.039)	(0.042)	(0.047)	(0.072)	(0.073)
	q-value	0.111	0.196	0.644	0.346	0.537	0.644
Obs.		1439	705	2144	1649	707	2065
Adjusted R2		0.606	0.494	0.169	0.843	0.563	0.885
Adjusted R2		0.606	0.502	0.175	0.843	0.566	0.888
Adjusted R2		0.616	0.495	0.177	0.844	0.570	0.887
Adjusted R2		0.607	0.493	0.163	0.843	0.557	0.885
Panel B: Peru							
Independent variables:							
Long-term memory	Coef.	0.072***	0.069**	0.054	-0.002	0.053	0.027
	s.e.	(0.018)	(0.028)	(0.032)	(0.015)	-0.037	(0.027)
	q-value	0.933	0.052	0.214	0.008	0.258	0.555
Inhibitory control	Coef.	0.031	0.026	-0.017	0.010	0.079	-0.021
	s.e.	(0.028)	(0.041)	(0.039)	(0.025)	-0.046	(0.046)
	q-value	0.768	0.614	0.718	0.679	0.183	0.718
Working memory	Coef.	0.093***	0.082**	0.042	0.024*	0.066**	-0.031
	s.e.	(0.014)	(0.037)	(0.030)	(0.013)	(0.028)	(0.035)
	q-value	0.144	0.081	0.322	0.045	0.067	0.529
Implicit learning	Coef.	0.013	0.070**	-0.006	-0.005	-0.070	-0.077
	s.e.	(0.024)	(0.029)	(0.032)	(0.012)	(0.056)	(0.059)
	q-value	0.747	0.063	0.854	0.899	0.307	0.389
Obs.		1692	676	2366	1718	669	2379
Adjusted R2		0.549	0.573	0.183	0.793	0.780	0.942
Adjusted R2		0.550	0.574	0.186	0.793	0.781	0.941
Adjusted R2		0.554	0.585	0.185	0.793	0.781	0.942
Adjusted R2		0.545	0.571	0.179	0.793	0.781	0.942

Note: Each coefficient in each cell comes from a different estimated model. Results from columns (1), (2), (4) and (5) are based on the value-added specification and follow Equation (2). Results from columns (1) and (4)—previously reported in Table 5—use data from the index children and (2) and (5) from the younger siblings. Results from columns (3) and (6) correspond to the value-added model with household fixed effects. ***p<0.01, **p<0.05, *p<0.10. q-values correct for multiple hypothesis testing.

Appendix A

Table A.1:
Associations between foundational cognitive skills at age 12 and test scores and highest grade attained at age 15 – All skills included simultaneously (Ordinary Least Squares)

Dependent variables:		**PPVT**	**Math**	**Reading comprehension**	**Highest grade attained**
		(1)	**(2)**	**(3)**	**(4)**
Panel A: Ethiopia					
Independent variables					
Long-term memory	Coef	0.115***	0.073**	0.107***	0.211***
	s.e.	(0.033)	(0.027)	(0.024)	(0.051)
Inhibitory control	Coef	0.034	0.099**	-0.007	0.270***
	s.e.	(0.041)	(0.037)	(0.036)	(0.067)
Working memory	Coef	0.115***	0.089***	0.067**	0.175***
	s.e.	(0.022)	(0.026)	(0.027)	(0.059)
Implicit learning	Coef	0.099	0.054	0.056	0.210*
	s.e.	(0.064)	(0.047)	(0.054)	(0.117)
Obs.		1469	1564	1541	1649
Adjusted R2		0.503	0.342	0.412	0.494
Panel B: Peru					
Independent variables					
Long-term memory	Coef	0.116***	0.095***	0.112***	0.021
	s.e.	(0.014)	(0.018)	(0.027)	(0.022)
Inhibitory control	Coef	0.044	0.097**	0.034	0.088*
	s.e.	(0.044)	(0.044)	(0.042)	(0.047)
Working memory	Coef	0.150***	0.142***	0.101***	0.119***
	s.e.	(0.023)	(0.029)	(0.028)	(0.033)
Implicit learning	Coef	-0.023	-0.050*	-0.036*	-0.030
	s.e.	(0.032)	(0.027)	(0.019)	(0.028)
Obs.		1695	1716	1684	1716
Adjusted R2		0.383	0.269	0.249	0.318

Note: In each panel, the four coefficients reported in each column come from the same estimated model. Each model controls for the baseline performance in all tasks, child´s sex, age in months, ethnicity, and number of years of pre-schooling; household wealth dummies (1 if in the middle tertile, 0 otherwise; 1 if in the top tertile, 0 otherwise), maternal level of schooling, area of residence, time of RACER administration (1 morning, 0 otherwise), date of RACER administration (1 if on a week day, 0 otherwise), and community-of-birth fixed effects. ***p<0.01, **p<0.05, *p<0.10.

Table A.2:
Ethiopia – Associations, one at a time, between foundational cognitive skills at age 12 and test scores at age 15 by language of administration

Dependent variables:		**PPVT – Amharic**	**PPVT - Tigrinya**	**PPVT – Oromifa**	**Reading comprehension - Amharic**	**Reading comprehension - Tigrinya**	**Reading comprehension - Oromifa**
		(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Panel A							
Independent variables:							
Long-term memory	Coef.	0.111**	0.135**	0.215*	0.148***	0.027	0.149*
	s.e.	(0.049)	(0.030)	(0.083)	(0.038)	(0.042)	(0.067)
Inhibitory control	Coef.	0.075	0.146	0.031	0.088*	0.034	-0.018
	s.e.	(0.073)	(0.083)	(0.088)	(0.046)	(0.075)	(0.091)
Working memory	Coef.	0.132***	0.119	0.180***	0.066*	0.134	0.060
	s.e.	(0.034)	(0.059)	(0.017)	(0.031)	(0.081)	(0.030)
Implicit learning	Coef.	0.086	0.317	0.097	0.038	0.212**	0.047
	s.e.	(0.086)	(0.159)	(0.119)	(0.058)	(0.060)	(0.240)
Obs.		777	354	320	680	335	274
Adjusted R2		0.570	0.301	0.252	0.385	0.116	0.119
Adjusted R2		0.576	0.308	0.212	0.375	0.138	0.097
Adjusted R2		0.588	0.340	0.267	0.380	0.156	0.111
Adjusted R2		0.561	0.319	0.211	0.367	0.132	0.103
Panel B							
Independent variables:							
Long-term memory	Coef.	0.076*	0.102*	0.188*	0.127***	0.001	0.135
	s.e.	(0.039)	(0.040)	(0.076)	(0.034)	(0.032)	(0.064)
Inhibitory control	Coef.	0.039	0.077	-0.006	0.057	0.004	-0.047
	s.e.	(0.072)	(0.059)	(0.082)	(0.045)	(0.105)	(0.049)
Working memory	Coef.	0.102***	0.085	0.163**	0.044	0.112	0.067*
	s.e.	(0.032)	(0.059)	(0.038)	(0.032)	(0.091)	(0.028)
Implicit learning	Coef.	0.039	0.236	0.075	0.006	0.131	0.050
	s.e.	(0.081)	(0.149)	(0.103)	(0.065)	(0.071)	(0.254)
Obs.		777	354	320	680	335	274
Adjusted R2		0.599	0.360	0.295	0.399	0.159	0.118

Note: ***p<0.01, **p<0.05, *p<0.10. In Panel A, each coefficient in each cell comes from a different estimated model. In Panel B, the four coefficients reported in each column come from the same estimated model. Each model was estimated only for those that answered the tests in Amharic, Tigrinya and Oromifa, respectively. Each model controls for the baseline performance in the associated task, child´s sex, age in months, and number of years of pre-schooling; household wealth dummies (1 if in the middle tertile, 0 otherwise; 1 if in the top tertile, 0 otherwise), maternal level of schooling, area of residence, time of RACER administration (1 morning, 0 otherwise), date of RACER administration (1 if on a week day, 0 otherwise), and community-of-birth fixed effects.

Table A.3:
Peru - Associations between foundational cognitive skills at age 12 and test scores at ages 15 among Spanish speakers

Dependent variables:		**PPVT – Spanish**	**Reading comprehension - Spanish**
		(1)	**(2)**
Panel A			
Independent variables			
Long-term memory	Coef.	0.148***	0.132***
	s.e.	(0.017)	(0.025)
Inhibitory control	Coef.	0.066	0.105**
	s.e.	(0.041)	(0.038)
Working memory	Coef.	0.192***	0.157***
	s.e.	(0.021)	(0.029)
Implicit learning	Coef.	0.043	0.050*
	s.e.	(0.030)	(0.024)
Obs.		1692	1738
Adjusted R2		0.318	0.260
Adjusted R2		0.330	0.278
Adjusted R2		0.347	0.280
Adjusted R2		0.302	0.250
Panel B			
Independent variables			
Long-term memory	Coef.	0.119***	0.113***
	s.e.	(0.014)	(0.028)
Inhibitory control	Coef.	0.047	0.033
	s.e.	(0.044)	(0.041)
Working memory	Coef.	0.149***	0.098***
	s.e.	(0.024)	(0.027)
Implicit learning	Coef.	-0.029	-0.039*
	s.e.	(0.032)	(0.020)
Obs.		1690	1678
Adjusted R2		0.373	0.239

Note: ***p<0.01, *p<0.05, *p<0.10. In Panel A, each coefficient in each cell comes from a different estimated model. In Panel B, the four coefficients reported in each column come from the same estimated model. Each model was estimated only for those that answered the tests in Spanish. Each model controls for the baseline performance in the associated task, child´s sex, age in months, and number of years of pre-schooling; household wealth dummies (1 if in the middle tertile, 0 otherwise; 1 if in the top tertile, 0 otherwise), maternal level of schooling, area of residence, time of RACER administration (1 morning, 0 otherwise), date of RACER administration (1 if on a week day, 0 otherwise), and community-of-birth fixed effects.

Table A.4:
Associations, one at a time, between foundational cognitive skills at age 12 and test scores and highest grade attained at age 15 (Value-added specifications) – Coefficients for the lagged test scores at age 12 (ρ)

Dependent variables:		**PPVT**	**Math**	**Reading comprehension**	**Highest grade attained**
		(1)	**(2)**	**(3)**	**(4)**
Panel A: Ethiopia					
Independent variables					
Long-term memory	Coef	0.543***	0.495***	0.455***	1,116***
	s.e.	(0.060)	(0.043)	(0.028)	(0.017)
Inhibitory control	Coef	0.539***	0.490***	0.462***	1.112***
	s.e.	(0,062)	(0.044)	(0.027)	(0.017)
Working memory	Coef	0.517***	0.486***	0.450***	1.106***
	s.e.	(0.061)	(0.042)	(0.026)	(0.016)
Implicit learning	Coef	0.550***	0.497***	0.461***	1.118***
	s.e.	(0.058)	(0.043)	(0.027)	(0.016)
Obs.		1439	1414	1368	1649
Adjusted R2		0.606	0.478	0.525	0.843
Adjusted R2		0.606	0.480	0.521	0.843
Adjusted R2		0.616	0.481	0.526	0.844
Adjusted R2		0.607	0.478	0.521	0.843
Panel B: Peru					
Independent variables					
Long-term memory	Coef	0.600***	0.602***	0.512***	1.068***
	s.e.	(0.045)	(0.032)	(0.023)	(0.026)
Inhibitory control	Coef	0.592***	0.599***	0.514***	1.061***
	s.e.	(0.044)	(0.032)	(0.023)	(0.026)
Working memory	Coef	0.580***	0.582***	0.498***	1.058***
	s.e.	(0.044)	(0.031)	(0.023)	(0.025)
Implicit learning	Coef	0.611***	0.608***	0.525***	1.068***
	s.e.	(0.045)	(0.031)	(0.023)	(0.025)
Obs.		1692	1710	1680	1718
Adjusted R2		0.549	0.482	0.400	0.793
Adjusted R2		0.550	0.481	0.397	0.793
Adjusted R2		0.554	0.486	0.408	0.793
Adjusted R2		0.545	0.482	0.395	0.793

Note: Each coefficient in each cell comes from a different estimated model. Unlike the previous table, each coefficient refers to the ρ value of the lagged test score obtained at age 12. Each model controls for the baseline performance in all (four) tasks, lagged test score, child´s sex, age in months, ethnicity, and number of years of pre-schooling; household wealth dummies (1 if in the middle tertile, 0 otherwise; 1 if in the top tertile, 0 otherwise), maternal level of schooling, area of residence, time of RACER administration (1 morning, 0 otherwise), date of RACER administration (1 if on a week day, 0 otherwise), and community-of-birth fixed effects. ***p<0.01, **p<0.05, *p<0.10.

Table A.5:
Associations between foundational cognitive skills at age 12 and test scores at age 15 – All skills included simultaneously (Value-Added Specification)

Dependent variables:		**PPVT**	**Math**	**Reading comprehension**
		(1)	**(2)**	**(3)**
Panel A: Ethiopia				
Independent variables				
Long-term memory	Coef.	0.038	0.017	0.068***
	s.e.	(0.027)	(0.029)	(0.019)
Inhibitory control	Coef.	0.014	0.068**	-0.056*
	s.e.	(0.036)	(0.025)	(0.029)
Working memory	Coef.	0.058***	0.045*	0.058**
	s.e.	(0.019)	(0.024)	(0.026)
Implicit learning	Coef.	0.099	-0.002	0.040
	s.e.	(0.057)	(0.035)	(0.044)
Obs.		1439	1414	1368
Adjusted R2		0.619	0.481	0.529
Panel B: Peru				
Independent variables				
Long-term memory		0.064***	0.045***	0.061***
		(0.017)	(0.015)	(0.021)
Inhibitory control		0.023	0.052*	-0.008
		(0.032)	(0.029)	(0.034)
Working memory		0.075***	0.051*	0.052*
		(0.017)	(0.025)	(0.025)
Implicit learning		-0.019	-0.046**	-0.028
		(0.027)	(0.021)	(0.022)
Obs.		1690	1708	1678
Adjusted R2		0.560	0.491	0.411

Note: In each panel, the four coefficients reported in each column come from the same estimated model. Each model controls for the baseline performance in all (four) tasks, lagged test score, child´s sex, age in months, ethnicity, and number of years of pre-schooling; household wealth dummies (1 if in the middle tertile, 0 otherwise; 1 if in the top tertile, 0 otherwise), maternal level of schooling, area of residence, time of RACER administration (1 morning, 0 otherwise), date of RACER administration (1 if on a week day, 0 otherwise), and community-of-birth fixed effects. ***p<0.01, **p<0.05, *p<0.10.

Table A.6:
Average performance by foundational cognitive skill among those not contacted for the phone survey versus those contacted during the phone survey

	N	Long-term memory	p-value	Inhibitory control	p-value	Working memory	p-value	Implicit learning	p-value
Panel b: Ethiopia									
Full sample (round 5)	1795	-0.13		-0.13		-0.15		-0.10	
Not contacted in Call 2	200	-0.20	0.294	-0.26	0.002	-0.35	0.002	-0.14	0.458
Contacted in Call 2	1595	-0.13		-0.11		-0.12		-0.09	
Panel A: Peru									
Full sample (round 5)	1841	0.13		0.13		0.15		0.09	
(a) Not contacted in call 5	212	0.03	0.122	-0.05	0.000	-0.11	0.000	0.06	0.618
(b) Contacted in call 5	1629	0.14		0.15		0.19		0.10	

Note: statistics correspond to the sample of index children that are part of the analytical sample. Group (a) are index children that were interviewed in the last in-person visit in 2016 (Round 5) but that were not contacted during the phone survey, whereas group (b) are those interviewed in Round 5 and also contacted for the phone survey

[i] A recent study that tackled this topic is Micalizzi et al. (2019). These authors focused on a group of four-year-old US children to assess how their performance varied according to SES. They concluded that children with higher SES showed higher levels of inhibitory control, which is associated with better school readiness.

[ii] One criterion from the beginning of the study was that country samples were to be pro-poor. This was implemented in different ways by each country team. The country team in Peru considered the universe of districts excluding the top 5% wealthiest districts, and randomly selected 20 districts from the remaining 95% (Escobal & Flores, 2008). In the case of Ethiopia, the country team used a purposive methodology. Clusters were chosen such that (i) poor areas were oversampled, (ii) the diversity across regions, ethnicity and location was captured; (iii) the costs of sampling were manageable (Outes-León & Sánchez, 2008).

[iii] In Ethiopia, elementary (primary) education officially starts at age 7 and lasts for 8 years, followed by 4 years of secondary education, whereas in Peru primary education is compulsory from age 6 and lasts for 6 years, followed by 5 years of secondary education.

[iv] Higher education is defined in Peru as enrolment in a technical institute or university, and for Ethiopia as any tertiary educational option that grants a Teacher´s certificate, a Diploma/Advanced diploma, a University degree, and a Masters or doctoral degree (it doesn´t include Technical and Vocation Education and Training).

[v] The YL wealth index takes values between zero and one, such that a larger value reflects a wealthier household. It is the simple average of a housing-quality index, an access-to-services index, and a consumer-durables index (Briones, 2017).

[vi] To see the relation explicitly between Equation (2) and the change in $EO_{i,c,t}$ note that Equation (2) can be modified to have the change in $EO_{i,c,t}$ as the explicit dependent variable by subtracting $EO_{i,c,t-1}$ from both sides.

[vii] We choose this specific procedure because it allows control the FWER (Familywise Error Control Rate), which in turn helps us keep differential effects by domain, that are of interest since the different tests might interact differently with each FCS

[viii] As mentioned, FCSs were standardized for the pooled sample of Ethiopia and Peru while the test scores (PPVT, math and reading comprehension) were standardized for the index children within each country for the main sample. The reasons behind this choice have been already explained. However, we note that our results are not sensitive to this standardization as we have also run our main results for Table 4 with the test scores standardized by age in years, resulting in virtually identical estimates. Moreover, in another approach, we also standardized the domain-specific test results for the pooled sample of children in both countries. In this case, most results remained significant, although the coefficients sometimes increased or decreased. However, for PPVT the results did change a bit more. We think this is because the adaptation of the PPVT test is the most sensitive to the cultural differences between countries (e.g., animals that were not common in Ethiopia or items which use was not common in Ethiopia were replaced).

[ix] Some coefficients become statistically insignificant, likely due to the smaller sample size when each country sample is partitioned. However, similar patterns are observed, except for reading comprehension in the Tigray region where the point estimates are very small.