School-entry socio-emotional skills and early grade literacy and mathematics achievement: Evidence from South Africa.

Authors Heleen Hofmeyr

Date July 2025

Working Paper 2025-07

PEP Working Paper Series

ISSN 2709-7331

Abstract

Existing research shows that socio-emotional skills are important determinants of major life outcomes. However, evidence of their early role in supporting the acquisition of foundational literacy and mathematics skills in low- and middle-income countries (LMICs) remains scarce. This paper investigates whether socio-emotional skills at school entry are associated with later literacy and mathematics achievement in South Africa. The results indicate that having higher socio-emotional skills at school entry is positively associated with Grade 2 literacy and mathematics scores, with stronger associations observed in better-resourced schools and among learners with higher cognitive skills at school entry. These findings underscore the potential value of integrating socio-emotional learning in early education to support academic development and reduce learning gaps in LMICs.

Keywords: socio-emotional skills; early academic achievement; foundational learning; South Africa; longitudinal data; educational inequality

Authors

Heleen Hofmeyr
Department of Economics, Stellenbosch University,
South Africa.
Email: heleenhofmeyr@sun.ac.za

Disclaimmer

This work was carried out with financial and scientific support from the Partnership for Economic Policy (PEP) and the Global Education Analytics Institute (GEAI) through funding provided by the Yidan Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of PEP or the GEAI.

Table Of Contents

I.Introduction

Millions of children in low- and middle-income countries (LMICs) fail to acquire foundational literacy and mathematics skills in the early grades (World Bank, 2018). In South Africa, for example, recent assessments show that the majority of children reach Grade 4 without acquiring the ability to read with understanding (Böhmer and Wills, 2023) or perform basic arithmetic operations (Venkat and Roberts, 2022). This learning crisis has far-reaching consequences, as early academic deficits tend to persist, and they widen learning gaps and limit later educational attainment and labour market prospects (Pritchett, 2013). Addressing these challenges requires a better understanding of the factors that support early learning and the policies that can enhance school readiness.

One important piece of this puzzle is the environments in which young children grow up. Many South African children are exposed to material deprivation, violence, food insecurity, and fractured family structures—all of which undermine their sense of safety, stability, and emotional development (Samuels, Slemming and Balton, 2012). These early adversities can have lasting effects on brain development and school readiness. While high-quality early childhood development (ECD) programmes have the potential to buffer children against these risks, access to them remains limited, particularly for children from low-income households (Moses and Van der Berg, 2023). As a result, many children enter school without the foundational cognitive and socio-emotional skills needed to succeed academically (Giese et al., 2022). In these contexts, socio-emotional competencies—such as the ability to manage emotions, avoid distractions, and form supportive relationships with peers and adults—may play a crucial role in helping young learners cope with adversity and remain engaged in learning. However, little empirical evidence exists on whether socio-emotional skills at school entry predict later academic achievement in South Africa or similar LMIC settings.

This paper addresses this gap in the literature by investigating whether socio-emotional skills at school entry predict early literacy and mathematics achievement in South Africa. More specifically, the research question is: To what extent do school-entry socio-emotional skills predict Grade 2 academic outcomes, even after controlling for baseline cognitive skills? Understanding this relationship is crucial, because socio-emotional skills have been linked to a range of positive life outcomes, including educational attainment, employment, and well-being (Heckman and Rubinstein, 2001; Roberts et al., 2007; Almlund et al., 2011; Danner, Lechner and Spengler, 2021). Determining their role in supporting foundational academic skills in LMICs could inform efforts to strengthen early learning interventions and improve long-term educational trajectories.

This study examines this issue using longitudinal data from the Roots and Shoots study, which follows a cohort of 400 South African learners from the start of Grade R (equivalent to kindergarten) through Grade 2. Ordinary least squares (OLS) regressions are estimated to assess the association between school-entry socio-emotional skills and later academic achievement. The analysis controls for a rich set of covariates, including baseline cognitive skills, socio-economic background, and indicators of school quality, to account for potential confounding factors. While the results do not support causal inference, a sensitivity analysis is conducted to assess how robust the observed associations are to omitted variable bias.

The findings show that socio-emotional skills at school entry are positively associated with later academic achievement: a 1-standard deviation (SD) increase in socio-emotional skills is associated with a 0.20-SD increase in Grade 2 mathematics achievement and a 0.14-SD increase in literacy achievement. These magnitudes are comparable to the short-term impacts of classroom quality improvements observed in early childhood settings (Chetty et al., 2011). They also align with meta-analytic findings regarding socio-emotional learning (SEL) interventions, which report average improvements in academic performance ranging from 0.18 to 0.46 SDs, with effects of 0.20 SDs or greater generally considered educationally meaningful (Durlak et al., 2011).

The sensitivity analysis results suggest that the association with mathematics achievement is robust to omitted variable bias under the assumption of equal selection on observables and unobservable, but the association with literacy is more sensitive to potential unobserved confounding. This pattern may indicate that socio-emotional skills are particularly important for acquiring mathematical competencies, which often require sustained effort and comfort with problem-solving in the face of uncertainty. There is also suggestive evidence of heterogeneity: the associations are stronger for boys, learners in better-quality schools, those from wealthier households, and those with higher baseline cognitive skills.

This study contributes to the growing literature on skill formation by providing new evidence of the role socio-emotional skills play in early academic achievement in Sub-Saharan Africa. While previous research has documented the long-term importance of socio-emotional skills for educational and labour market outcomes (Heckman and Kautz, 2012; Cunha et al., 2006; Danon et al., 2024; Sorrenti et al., 2025), the influence these skills have during the early years of schooling remains underexplored—particularly in LMICs (Wolf and McCoy, 2019; Zehner et al., 2024). This study highlights the contextual factors that moderate these associations—such as school quality and socio-economic background—and offers actionable insights for policymakers seeking to improve foundational learning. The findings also support the case for integrated early childhood programmes that combine socio-emotional and cognitive development in line with recent experimental evidence from LMICs (Lopez Garcia et al., 2023; Díaz et al., 2023). Taken together, the results highlight the potential of early investments in SEL to narrow educational inequalities and improve long-run outcomes (García et al., 2020).

The remainder of this paper is structured as follows. Details about the South African context are presented in Section 2. Section 3 outlines the data and key variables used, and presents sample statistics as well as a brief discussion of sample attrition. Section 4 sets out the empirical framework used. Section 5 presents the main results, including robustness checks and heterogeneity analyses, and a discussion of potential theoretical mechanisms. Section 6 concludes with a summary of the key findings and their implications for policy and future research.

II.The South African context

Formal schooling in South Africa begins in Grade 1 at age six, but nearly all government primary schools now include a preschool year—Grade R—for five-year-olds (Pretorius, Rastle and Mtsatse, 2022). Learners are typically taught in their home language in Grades R to 3, after which instruction shifts to English or Afrikaans. The academic year begins in January and is broken up into four terms, each of which lasts roughly three months.

Many South African children grow up in contexts marked by poverty, violence, food insecurity, and fractured family structures (Samuels, Slemming and Balton, 2012). Although

ECD programmes can buffer against early life adversity, access to them remains limited, particularly for children from low-income households (Moses and Van der Berg, 2023). Many children thus enter Grade R—their first structured learning environment—without the foundational skills needed to succeed in school (Giese et al., 2022; Richter and Samuels, 2018). While local researchers and policymakers are starting to pay attention to the cognitive dimension of school readiness, socio-emotional skills have remained largely overlooked.

In addition, it is not known how school quality interacts with the extent to which socio-emotional skills support early learning. This is an important question to consider in the South African context given the sharp structural inequalities that exist in the education system. Despite extensive post-apartheid reforms, learners in historically black and coloured communities—which are still shaped by spatial and economic exclusion—attend predominantly no-fee public schools. These schools, which serve around 80% of learners, are under-resourced, with large class sizes, insufficient materials, and poor infrastructure (Spaull and Taylor, 2022). In contrast, fee-charging schools—often formerly white schools—offer significantly better learning environments and attract more affluent learners. Outcomes differ greatly by school fee status: in 2021, 90% of children in no-fee schools could not read with comprehension by Grade 4, compared with just 10% in fee-charging schools (Böhmer and Wills, 2023). As a result, school fee status serves as a useful and policy-relevant proxy for school quality in South Africa (Yamauchi, 2011) and is used in this study to control for differences in school quality across learning environments.

III. Data

This section presents the data used in this study. First, the study design and sample selection are detailed, then the data collection instruments are discussed, and lastly, summary statistics and a brief discussion about attrition are provided.

3.1 Study design and sample selection

This study uses data from the Roots and Shoots project, a longitudinal study that was designed to examine how socio-economic and linguistic inequalities in school readiness contribute to South Africa's persistent learning disparities (Hofmeyr, Ardington and Spaull, 2022). The project followed a cohort of children from the start of Grade R (kindergarten) in 2022 (Wave 1) to the third term of Grade 2 in 2024 (Wave 3) across 75 primary schools in the Western Cape province. A total of 556 children were assessed in Wave 1, and 400 learners (72%) were successfully re-assessed in Wave 3 with complete data for all the outcome and control variables.

The sample includes schools in five education districts[1] in urban and peri-urban areas in and around Cape Town. The majority of schools (73%) serve historically disadvantaged communities and do not charge school fees. Since the Roots and Shoots project sought to investigate socio-economic disparities in school readiness, its sample also included fee-charging schools, which are mostly located in suburban areas. Fifty of the schools were Afrikaans-medium, and twenty-five, isiXhosa-medium, which represent two of the largest language groups in the Western Cape province. Although the sample is not nationally or provincially representative, it captures meaningful variation in socio-economic status (SES), school quality, linguistic background, and learner performance in the early grades.

Roots and Shoots was part of a broader impact evaluation of a literacy intervention implemented by the NGO Funda Wande. The final sample includes learners from both the treatment and control schools in the original evaluation as well as additional isiXhosa-medium schools selected to broaden linguistic representation[2]. Figure 1 presents a Venn diagram illustrating the overlap between the schools included in the Roots and Shoots study and those participating in the Funda Wande evaluation. Fifty Afrikaans schools were part of both studies,

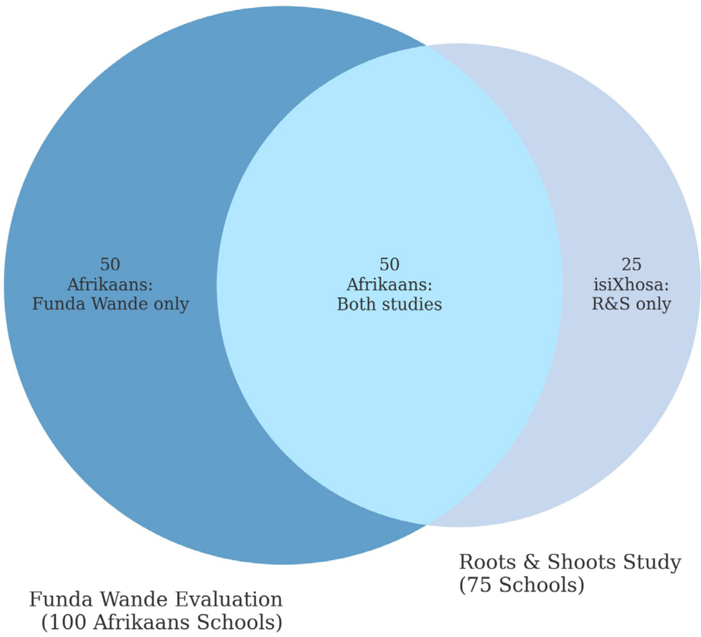
[1] The districts are Metro Central, Metro East, Metro North, Metro South and Cape Winelands.

[2] The Afrikaans schools were selected from the 100 schools in the Funda Wande evaluation, with 25 treatment schools and 25 matched control schools drawn from the districts closest to Cape Town. Stratification by district and school fee status guided selection. The isiXhosa schools were drawn from the same districts but stratified by performance quintiles using systemic test scores. See the Appendix for full sampling details.

50 Afrikaans schools were unique to the Funda Wande evaluation, and the 25 isiXhosa schools were unique to Roots and Shoots.

One Grade R class was randomly selected from each school, and eight learners were randomly sampled from each class to arrive at a planned sample of 600 learners. Due to caregiver withdrawal (N = 11) and unsuitable testing conditions on the day of the assessment[3] (N = 33), the final baseline sample consisted of 556 learners. Children were assessed in their home language (Afrikaans or isiXhosa) by trained enumerators using standardized assessments. Demographic information was collected through learner interviews and school records. A total of 410 learners from the original sample were assessed again in Wave 3 (73.7%). The analysis is restricted to the 400 learners with complete data for all the variables used in the analysis.

Figure 1: Overlap between schools in the Funda Wande evaluation and the Roots and Shoots study



Notes: This Venn diagram illustrates the overlap in schools sampled between the Funda Wande evaluation and the Roots & Shoots (R&S) study. The Funda Wande evaluation included 100 Afrikaans-speaking schools, 50 of which also participated in the Roots & Shoots study. The Roots & Shoots sample comprised 75 schools: 50 Afrikaans-speaking schools that also participated in the Funda Wande

[3] Conditions were deemed unsuitable if the enumerator could not find a quiet space to assess the learner one-on-one or the assessment could not continue due to interruptions.

evaluation, and an additional 25 isiXhosa-speaking schools that were unique to Roots & Shoots. Data were collected for both studies concurrently.

3.2 Survey instruments

3.2.1 School-entry cognitive and socio-emotional skills

School readiness was assessed in Wave 1 using the Early Learning Outcomes Measure (ELOM) 4&5. ELOM 4&5 is a locally developed tool (Dawes et al., 2016) that makes use of one-on-one direct observation of a number of tasks administered by a trained enumerator to assess development in five domains of school readiness, namely gross motor development, fine motor coordination and visual integration, cognition and executive function, emergent numeracy and mathematics, and emergent literacy and language. Learners' scores for the latter three domains—cognition and executive function, emergent numeracy and mathematics, and emergent literacy and language—are used to measure their school-entry cognitive skills. A single school-entry cognitive skills score was calculated for each learner by converting their scores for each of the three domains into percentages (by dividing the number of correct items by the total number of items making up that domain) and then calculating an unweighted average across the three domains. Scores were standardized to have a mean of zero and an SD of one.

Socio-emotional skills were measured in Wave 1 using the ELOM Social and Emotional Functioning scale (Dawes et al., 2016). This scale is administered to teachers and is intended to capture two aspects of students' social and emotional development, namely their social relations with peers and adults, and their emotional readiness for school. For the social relations with peers and adults subscale, teachers were asked to rate students on six items based on how often they exhibit certain behaviours ("None of the time", "A little of the time", "Most of the time" and "All of the time"). The items making up the social relations subscale were drawn from the Child Trends Teacher Rating (Child Trends, 2014) and the California Desired Results Developmental Profile (California Department of Education, 2008, 2010). For the emotional readiness for school subscale, teachers rated students on six statements (with the options "Not true", "Sometimes true" and "Often true"). These items were selected from the South African Child Assessment Scales (SACAS) questionnaire, which is based on the

Achenbach Child Behaviour Checklist (Barbarin and Richter, 2001). The items comprising these two subscales can be found in the Appendix. Teachers' responses to the 12 items in the ELOM Social and Emotional Functioning scale were combined by computing the sum of the answers for each subscale[4] and then calculating a weighted average[5] score across the two subscales. Scores were standardized to have a mean of zero and an SD of one.

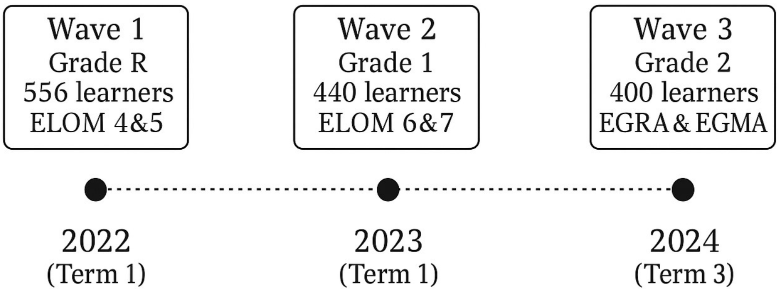
3.2.2 Achievement in literacy and mathematics

Literacy skills were assessed by enumerators during one-on-one observation using the Early Grade Reading Assessment (EGRA). This tool was developed by Research Triangle International (RTI) and adapted for use in Afrikaans (Ardington, Mohohlwane and Barends, 2022) and isiXhosa (Ardington et al., 2020) by local experts. The adapted EGRA contains seven tasks that test learners on letter-sound recognition, phonemic awareness, syllable reading, word reading, passage reading (oral reading fluency), reading comprehension and listening comprehension. Mathematics skills were assessed in a group setting using the written Early Grade Mathematics Assessment (EGMA), which was also developed by RTI and translated into Afrikaans and isiXhosa by local experts. The EGMA contains seven tasks that test learners on number concept, addition and subtraction of single-digit and double-digit numbers, pattern recognition and word problems. As is the case for literacy, each task consists of a number of items. For example, learners are given 20 single-digit subtraction problems and 5 word problems. A three-step process was used to construct overall scores for literacy and mathematics. First, learners' scores for each task were standardized. The standardized task scores were then averaged to obtain a composite score for each subject. Finally, the composite scores were standardized to facilitate interpretation and comparability. This process ensured that each task contributes equally to the final subject score regardless of its original scale. Figure 2 presents the Roots and Shoots study's data collection timeline.

[4] For the social relations with peers and adults subscale, a response of "None of the time" was coded as 0, "A little of the time" was coded as 1, "Most of the time" was coded as 2 and "All of the time" was coded as 3. For the emotional readiness for school subscale, "Not true" was coded as 0, "Sometimes true" was coded as 1 and "Often true" was coded as 2.

[5] So that the scores for each subscale contribute equally to the total score.

Figure 2: Longitudinal data collection timeline: Roots and Shoots study



Notes: This figure shows the three waves of data collection that were performed for the Roots & Shoots study. Wave 1 took place during Term 1 of Grade R in 2022 and included 556 learners assessed using the ELOM 4&5 instrument. Wave 2 was conducted during Term 1 of Grade 1 in 2023, with 440 learners assessed using the ELOM 6&7 instrument. Wave 3 took place during Term 3 of Grade 2 in 2024, with 400 learners assessed using the EGRA and the EGMA. The timeline reflects the longitudinal tracking of learners across these three time points.

3.2.3 Learner and school characteristics

In addition to learners' cognitive skill level at school entry, several learner and school characteristics were collected for use in the analysis. Learners' age and gender were obtained from school records in Wave 1 and confirmed by enumerators when assessing children. Learners' grade level in Wave 3 was also obtained from school records and verified during assessments. Learners' SES was captured using two indicators: an asset index score and a binary indicator of household social grant receipt. Asset index scores were constructed using principal component analysis (PCA) based on learners' responses to the questions that were asked about 13 household assets[6] during the Wave 3 assessments, with pictures used to aid comprehension. Information on household social grant receipt was collected from

[6] The assets included in the questionnaire were: cell phone, computer/tablet, electricity, refrigerator, washing machine, television, radio, stove, bicycle, car, toilet inside the house and tap water.

teachers. The school fee status (no-fee or fee-charging) and the language of instruction were obtained from administrative records provided by the Western Cape Education Department.

3.3 Summary statistics and attrition

Summary characteristics are presented in Table 1 for the 400 learners who were assessed in both Wave 1 and Wave 3. The mean age for the sample is 5.45 years, and 49% (N = 195) of learners in the sample were male. In Wave 3, 88% of the learners assessed were in Grade 2. By construction, the asset index has a mean value of zero and an SD of one. Fifty-five percent of learners came from households that received a social grant, which is an indication of socio-economic disadvantage. Just over a quarter of learners (29%, N = 115) attended fee-charging schools, and 33% of learners were in isiXhosa-medium schools. All the performance variables are approximately standard normal, with means close to zero and SDs near one, as expected by construction. The range of values, however, points to meaningful differences across learners. The socio-emotional skills at school entry variable has a wider left tail (min = -3.15) than the cognitive skills at school entry variable (min = -2.55), which suggests greater variability in children's socio-emotional skills than their cognitive skills at school entry. Academic achievement in Grade 2 also varies widely. The literacy and mathematics z-scores each span more than 4 SDs, which highlights substantial learning inequalities by the middle of Grade 2.

Table 1: Summary statistics

	Mean	SD	Min.	Max.
Learner characteristics				
Age (Wave 1)	5.45	0.22	4.25	7.75
Male	48.75%			
Learner is in Grade 2 (Wave 3)	87.75%			
Household characteristics				
Asset index	0.00	1.00	-4.47	0.96
Household receives social grants	54.75%			
School characteristics				
School charges fees	28.75%			
Language of instruction is isiXhosa	33.00%			
Performance				
Cognitive skills at school entry (Wave 1)	0.03	1.01	-2.55	2.25
Socio-emotional skills at school entry (Wave 1)	-0.01	1.01	-3.15	1.64
Literacy (Wave 3)	0.00	1.00	-2.48	1.81
Mathematics (Wave 3)	0.00	1.00	-2.58	2.24

Source: Roots and Shoots study. Notes: Learner age is measured in years. The asset index scores were calculated using PCA for the 13 items learners indicated having or not in their homes. Scores were standardized to have a mean of 0 and an SD of 1. The reference categories for the school characteristics were: school does not charge fees, and language of instruction is Afrikaans. Performance scores are reported in SDs.

To assess potential attrition bias, the baseline characteristics of learners who were retained and assessed in Wave 3 are compared with those of learners who attrited between Waves 1 and 3. Table 2 reports mean values by group and p-values across the groups for key variables. Most of the characteristics are balanced; however, attrited learners were slightly younger at baseline ($p = 0.040$) and significantly less likely to report household social grant receipt ($p = 0.000$), which suggests some non-random attrition by SES. Crucially, there are no significant differences in baseline socio-emotional skills ($p = 0.649$) or cognitive skills ($p = 0.266$), which are the primary predictors of interest. This supports the internal validity of the analysis, as attrition does not appear to be systematically related to key explanatory variables.

Table 2: Balance tests between retained and attrited learners

Baseline variable	Mean (Retained)	Mean (Attrited)	Difference	p-value
Socio-emotional skills	-0.012	0.031	-0.043	0.649
Cognitive skills	0.028	-0.078	0.105	0.266
Female	0.514	0.465	0.049	0.299
Age (years)	5.455	5.409	0.544	0.040
Social grant recipient	0.549	0.026	0.523	0.000
Afrikaans	0.668	0.606	0.062	0.170
Fee-charging school	0.287	0.252	0.035	0.407

Notes: The table shows how the baseline characteristics of retained and attrited learners differ using two-tailed t-tests. The socio-emotional skills and cognitive skills scores were standardized to have a mean of 0 and an SD of 1. The results show that the retained learners were significantly older and more likely to be social grant recipients, which indicates lower SES.

IV.Empirical framework

To examine whether school-entry socio-emotional skills predict later academic achievement, Grade 2 literacy and mathematics scores are regressed on baseline socio-emotional skill levels. The models control for school-entry cognitive skills as well as a rich set of learner, household, and school characteristics to account for observable differences that may confound the relationship. Socio-emotional skills at school entry are shaped by early life experiences, home environments, and individual traits, many of which are difficult to observe and may also influence academic outcomes. While controlling for cognitive skills and other observables mitigates some bias, unobserved confounding remains a concern. The results are therefore interpreted as correlational, though they do offer suggestive evidence of the role socio-emotional skills potentially play in shaping early learning trajectories.

To examine the effect socio-emotional skills have on foundational literacy and mathematics skills, the following model inspired by Todd and Wolpin (2003) was estimated:

$$Y_{it} = \alpha + \beta X_{it-1} + \gamma \mathbf{Z_i} + \epsilon_i \quad (1)$$

where Y_{it} represents the Grade 2 literacy or mathematics score of student *i*, X_{it-1} represents student *i*'s socio-emotional skills score at school entry, and $\mathbf{Z}_i$ is a vector of controls (cognitive skills at school entry, gender, age, social grant receipt, school fee status, school language of instruction and school treatment status[7]). The main coefficient of interest is β. The inclusion of learners' school-entry cognitive skills is particularly important. Because cognitive ability is likely correlated with both socio-emotional skills and later academic outcomes, controlling for it helps to reduce omitted variable bias and better isolate the relationship of interest.

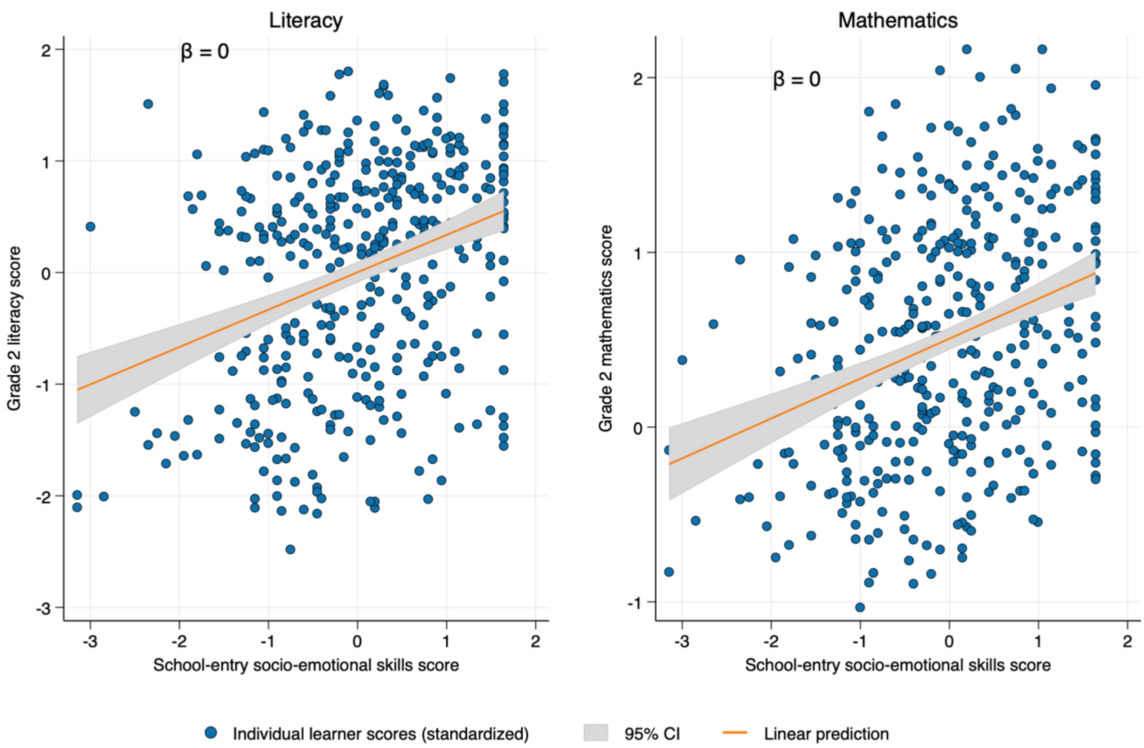
The OLS regression analysis was conducted in a sequential manner, with the association between socio-emotional skills and later achievement estimated first in the absence of controls (i.e., Equation (1) without $\mathbf{Z}_i$) and then with controls. Given the relatively limited set of controls that relate to school quality (school fee status, language of instruction and treatment status), the models were also estimated with school fixed effects (FEs). This sequential strategy makes it possible to evaluate the stability of the coefficients on socio-emotional skills and the R^2 values in the absence and presence of controls, which provides an indication of the extent to which the controls explain the raw differences in literacy and mathematics scores of learners with different levels of socio-emotional skills (Altonji, Elder and Taber, 2005).

[7] Treatment status refers to whether a school was a treatment or control school in the Funda Wande evaluation.

V.Results

The unconditional association between school-entry socio-emotional skills and later literacy and mathematics achievement is investigated first before proceeding with the estimation of these variables' conditional association. Figure 3 plots the Grade 2 literacy and mathematics test scores associated with each learner's school-entry socio-emotional skills score (indicated by the blue dots). Linear prediction is used to estimate the association between these variables (indicated by the orange line). The 95% confidence intervals around the fitted values are indicated by the grey shading. The slope of each fitted line corresponds with β in Equation (1) above (without Z_i). Figure 3 shows that both literacy and mathematics scores in Grade 2 are highly correlated with school-entry socio-emotional skills: A 1-SD increase in socio-emotional skills is associated with a 0.33-SD increase in literacy scores and a 0.34-SD increase in mathematics scores. Next, the regression analysis determines whether this strong association holds in a multivariate context.

Figure 3: Unconditional association between school-entry socio-emotional skills and Grade 2 literacy and mathematics scores



Notes: The figure shows the associations between school-entry socio-emotional skills and Grade 2 literacy (left panel) and mathematics (right panel) scores. Scores were standardized to have a mean of 0 and an SD of 1. Each dot on the scatter plot represents the combination of an individual learner's school-entry socio-emotional skills score and Grade 2 literacy/mathematics score. The orange lines represent linear predictions of the association between school-entry socio-emotional skills and Grade 2 literacy/mathematics scores. The areas shaded grey indicate the 95% confidence intervals around the fitted values.

5.1 Main results

It is now time to study how predictive school-entry socio-emotional skills are of early grade literacy and mathematics achievement. More specifically, the conditional association between school-entry socio-emotional skills and later academic achievement is estimated while controlling for school-entry cognitive skills and additional controls. The OLS regression estimation results are presented in Table 3. Only the coefficients on school-entry socio-emotional skills are presented in the table since this is the main effect of interest; however, the coefficients on the full set of covariates are presented in Table A1 in the Appendix. Since standardized variables were used to measure cognitive, socio-emotional, literacy and mathematics skills, the coefficients are reported in SDs.

The coefficients presented in the first two columns of Table 3 match the slopes of the fitted lines in Figure 3, since Models (1) and (2) estimate the association between socio-emotional skills and academic achievement in the absence of controls. When controls are added, these coefficients are reduced substantially, from 0.334 to 0.141 for literacy and from 0.341 to 0.198 for mathematics. The R^2 values also increase when controls are added, from 0.113 to 0.313 in the literacy model and from 0.111 to 0.331 in the mathematics model. These results indicate that the controls (particularly school-entry cognitive skills and school fee status—see Table A1 in the Appendix) are powerful predictors of later literacy and mathematics achievement and lead to a relatively large change in the estimated effect of school-entry socio-emotional skills. However, it is notable that a 1-SD increase in socio-emotional skills is associated with a 0.141-SD increase in literacy scores and a 0.198-SD increase in mathematics scores even when controls are factored in. This is particularly noteworthy since the set of controls includes school-entry cognitive skills, so these values indicate that school-entry socio-emotional skills are associated with later literacy and mathematics achievement even when conditioning on school-entry cognitive skills.

The results in Table 3 show that the estimated coefficients on socio-emotional skills increase when school FEs are included: from 0.141 to 0.248 SDs for literacy and from 0.196 to 0.260 SDs for mathematics. This pattern suggests that the predictive power of socio-emotional skills is stronger within schools than across them. One interpretation is that unobserved school-level factors—such as average school quality—mask the within-school relationship when not accounted for. Including school FEs absorbs these school-level differences, thereby isolating variation across learners within the same school. While the inclusion of school FEs lowers the R^2 values (from 0.318 to 0.244 in the literacy model and from 0.335 to 0.305 in the mathematics model), this is expected, as these models no longer capture between-school variance in achievement. Nonetheless, socio-emotional skills remain strongly predictive of later outcomes, even when school-level heterogeneity is controlled for.

To further investigate the role of school-level heterogeneity, the extent to which variation in academic achievement occurs between versus within schools is also assessed. Following Rodriguez-Segura and Tierney (2024), the total variance in literacy and mathematics outcomes was decomposed into within- and between-school components using the intraclass correlation coefficient (ICC). The ICC captures the proportion of total variance attributable to between-school differences and is calculated as follows:

$$\text{ICC} = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2} \tag{2}$$

where σ_B^2 denotes the variance between schools (i.e., the variance of school means) and σ_W^2 denotes the average variance within schools. This decomposition makes it possible to quantify the extent to which academic achievement is clustered at the school level. The ICCs for literacy and mathematics are 0.22 and 0.21, respectively, which indicates that only around one-fifth of the variation in learner achievement occurs between schools, while the majority—approximately 78% to 79%—occurs within schools. This suggests that although school-level factors matter, much of the variability in performance arises among learners attending the same school. This finding reinforces the relevance of individual-level characteristics—

including socio-emotional skills—which are notably associated with both literacy and mathematics achievement in the analysis.

Table 3: OLS results for the effect school-entry socio-emotional skills have on later literacy and mathematics scores

	(1)	(2)	(3)	(4)	(5)	(6)
	Literacy	**Math**	**Literacy**	**Math**	**Literacy**	**Math**
Socio-emotional skills	0.334***	0.341***	0.141**	0.198***	0.248***	0.260***
	(0.065)	(0.060)	(0.062)	(0.055)	(0.068)	(0.062)
Constant	0.004	0.004	-2.688**	-0.632	-0.327	0.166
	(0.071)	(0.072)	(1.116)	(0.908)	(1.159)	(1.095)
Controls	No	No	Yes	Yes	Yes	Yes
School FEs	No	No	No	No	Yes	Yes
N	400	400	400	400	400	400
R^2	0.113	0.118	0.318	0.335	0.244	0.305

Notes: School-entry socio-emotional skills and Grade 2 literacy and mathematics scores were standardized to have a mean of 0 and a standard deviation of 1. Standard errors are reported in parentheses. Asterisks indicate statistical significance at * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Models 1 and 2 estimate the effect in the absence of controls. Models 3 and 4 control for school-entry cognitive skills, gender, age, social grant receipt, asset index scores, school fee status, school language of instruction and school treatment status. Models 5 and 6 control for school-entry cognitive skills, gender, age, social grant receipt and asset index scores.

5.2 Robustness checks

5.2.1 Accounting for potential attrition bias

To assess whether sample attrition between Wave 1 and Wave 3 could bias the results, the main models were re-estimated using inverse probability weighting (IPW). Although balance tests showed no significant differences in key baseline characteristics between retained and attrited learners, IPW adjusts for potential selection bias based on the joint distribution of observed variables, which provides a more sensitive robustness check. First, the probability of being retained in the sample was estimated using a logit model that regresses retention status on baseline learner and school characteristics[8]. Then, the inverse probability weights were calculated as the reciprocal of the predicted retention probabilities. Learners who were less likely to be retained are weighted more heavily in the weighted regressions. The results, which are reported in Table A2 in the Appendix, show that while the coefficients on socio-emotional skills remain statistically significant after applying IPW, the estimated magnitudes decrease notably, particularly for mathematics achievement. The coefficient on literacy achievement declines modestly from 0.14 to 0.12 SDs, while the coefficient on mathematics achievement declines more substantially from 0.20 to 0.12 SDs. This attenuation suggests that sample attrition was not entirely random with respect to socio-emotional skills and later academic outcomes. However, the results still indicate a positive and statistically significant association between school-entry socio-emotional skills and Grade 2 literacy and mathematics achievement after adjusting for this selection bias. This suggests that while attrition led to some upward bias in the main estimates—particularly for mathematics achievement—the core finding that socio-emotional skill level at school entry predicts later academic outcomes remains robust and statistically significant after accounting for observable selection bias associated with sample attrition.

5.2.2 Sensitivity to omitted variable bias

Beyond potential attrition bias, omitted variable bias within the retained sample may also be a concern since there may be unobserved learner/household characteristics that correlate with both socio-emotional skills at school entry and later academic achievement and would confound the estimated effect school-entry socio-emotional skills have on later achievement. For example, parents who dedicated more time to activities that promoted their children's socio-emotional development prior to school entry likely went on to dedicate time to their children's academic achievement by helping with homework.

[8] The following baseline covariates were used to estimate the retention model: school-entry socio-emotional and cognitive skills, learner age, gender, SES (asset index and grant receipt), school fee status, school language of instruction and school treatment status.

The extent to which the results are sensitive to omitted variable bias was investigated using the method that Oster (2019) proposed to assess the stability of the coefficients on socio-emotional skills in the presence of selection on unobservables. This method, which builds on the work of Altonji, Elder and Taber (2005), assumes that the degree of selection on unobservables can be determined from the degree of selection on observables. Since the results presented in Table 3 show that the magnitude of the coefficients on school-entry socio-emotional skills changes quite substantially when controls are added to the regression, it is clear that the effect is at least partly driven by selection on observables. If it is assumed, as Altonji, Elder and Taber (2005) argue one can, that the degree of selection on unobservables is a function of the degree of selection on observables, then the fact that there is selection on observables implies there is likely selection on unobservables as well. A sensitivity analysis was conducted to assess how robust the results are to omitted variable bias. The results are presented below.

Oster (2019)'s method requires making assumptions about: (i) the relative degree of selection on observables and unobservables, which is denoted by δ, and (ii) the R^2 from a hypothetical regression of the outcome on both observed and unobserved controls, which is denoted by R_{max}. If later achievement could be fully explained by the set of controls (both observed and unobserved), then R_{max} would equal 1. However, given the likelihood of measurement error in both the outcome and the covariates, a more realistic value of R_{max} is some function of the estimated R^2. In accordance with Oster (2019), it is assumed that R_{max} = 1.3 × R^2_{full}, where R^2_{full} is the R^2 from the regression that includes the full set of observable controls. The rationale behind this assumption is that if all relevant unobserved factors were accessible, they would likely increase the explanatory power of the model, but not to an extreme degree. Oster derived this 1.3 multiplier from an empirical analysis of randomized controlled trials that showed the total explainable variation in an outcome is typically no more than 30% greater than what is explained by observables alone. This assumption strikes a balance between acknowledging the existence of omitted variables and recognizing that a substantial portion of outcome variation is due to idiosyncratic factors, measurement error, or randomness, which no set of controls can fully capture.

The bias-adjusted estimate is calculated as follows:

$$\hat{\beta}^* = \hat{\beta}_{full} - \delta \cdot \left(\hat{\beta}_{full} - \hat{\beta}_{restricted}\right) \cdot \frac{R_{\max} - R^2_{full}}{R^2_{full} - R^2_{restricted}} \tag{3}$$

where $\hat{\beta}_{restricted}$ and $R^2_{restricted}$ are obtained from a regression without controls (Models 1 and 2 in Table 3), and $\hat{\beta}_{full}$ and R^2_{full} come from a regression with the full set of controls (Models 3–6 in Table 3). The variable δ captures the assumed relative strength of selection on unobservables compared with selection on observables.

The bounds on the coefficient of interest are estimated for different values of δ, specifically $\delta = 0$, which corresponds to the original OLS estimates (assuming no selection on unobservables), and $\delta = 1$, which assumes that the unobservables are as strongly correlated with the outcome as the observables. The results are presented in Table 4.

The results show that under these assumptions, the lower bound of the estimated effect socio-emotional skills have on mathematics achievement (reported in the third column of Table 4) remains positive, which suggests that the result is robust to omitted variable bias. In contrast, the lower bound of the effect socio-emotional skills have on literacy achievement becomes negative in the model that includes school FEs. This indicates that the association between socio-emotional skills and literacy achievement could be driven by unobserved confounders and may not be robust when it is assumed that selection on unobservables is as strong as selection on observables. In other words, the positive association observed in Models 3–6 could plausibly be explained away by omitted variables that are similarly predictive as the included controls.

Since the regressions include a rich set of covariates—including school-entry cognitive skills, school language of instruction, SES, and school quality—it may be unlikely that the unobservables would explain substantially more variation in literacy achievement or exert stronger effects than the observed controls. Nonetheless, the sensitivity of the results for literacy achievement suggests that this association should be interpreted with caution. The more robust findings for mathematics achievement provide greater confidence that socio-emotional skills play a meaningful role in supporting early mathematics development, even in the presence of potential omitted variable bias.

Table 4: Estimated bounds of coefficients on school-entry socio-emotional skills in the presence of omitted variable bias

	Coefficient on socio-emotional skills	Bounds	School FEs
Literacy	0.141 (0.062)	[0.037; 0.141]	No
Math	0.198 (0.055)	[0.115; 0.198]	No
Literacy	0.248 (0.067)	[–0.093; 0.248]	Yes
Math	0.260 (0.060)	[0.158; 0.260]	Yes

Notes: This table shows the results of estimating the bounds (Oster, 2019) of the coefficients on school-entry socio-emotional skills obtained from the main analysis (shown in the second column). Standard errors calculated at the school level are reported in parentheses. The bounds are calculated using $R_{max} = 1.3R^2$, $\delta = 0$ (in line with the original estimates), and $\delta = 1$ as the lower bound, which corresponds with the assumption of equal selection between observed and unobserved variables, as suggested by Oster (2019). The results show that the models predicting mathematics achievement are robust to selection on unobservables, whereas the models predicting literacy achievement are not.

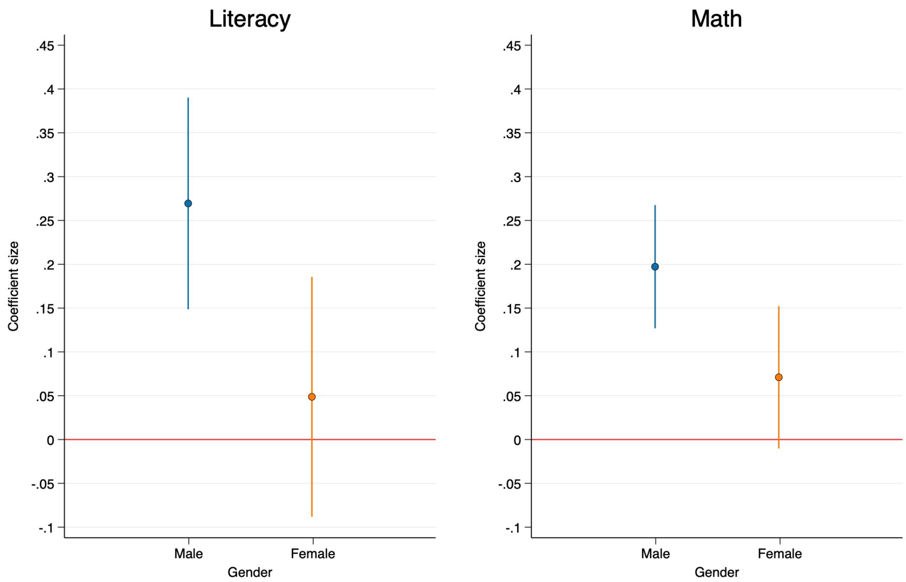
5.3 Heterogeneous effects

The results presented thus far have assumed that school-entry socio-emotional skills have the same conditional association with later literacy and mathematics achievement for all learners. However, there is reason to believe the association may differ along a number of dimensions, such as gender, school quality, learner SES and level of cognitive skills at school entry. This question was investigated by estimating Equation (1) separately for boys and girls, learners in no-fee and fee-charging schools, those in the bottom and top halves of the SES distribution, and those in the bottom, middle and top thirds of the distribution of cognitive skills scores at school entry. The results point to significant differences in the magnitude of the association between school-entry socio-emotional skills and later academic achievement along all of the dimensions considered. These results are discussed in more detail below.

5.3.1 Gender

Other studies that have made use of longitudinal data to estimate the effect socio-emotional skills have on academic achievement have found evidence of heterogeneous effects by gender (Duncan et al., 2007; Li-Grining et al., 2010). This is a particularly important question in the South African context, as girls exhibit a large and persistent educational advantage in South Africa (Spaull and Makaluza, 2019). Figure 4 shows the coefficients obtained when estimating the conditional association between school-entry socio-emotional skills and later literacy and mathematics achievement separately for boys and girls. The figure shows the magnitude of the coefficients on socio-emotional skills clearly differs by gender, with the coefficient estimated for boys being much larger than the one estimated for girls in the case of both literacy and mathematics achievement. The coefficient on socio-emotional skills is statistically significant for boys for both subjects, but not for girls for either subject. However, the absence of statistical significance for girls may reflect limited statistical power rather than a true null effect. These findings should therefore be interpreted as suggestive of potential gender differences in the returns on socio-emotional skills.

Figure 4: Conditional associations between school-entry socio-emotional skills and later academic achievement, by subject and gender

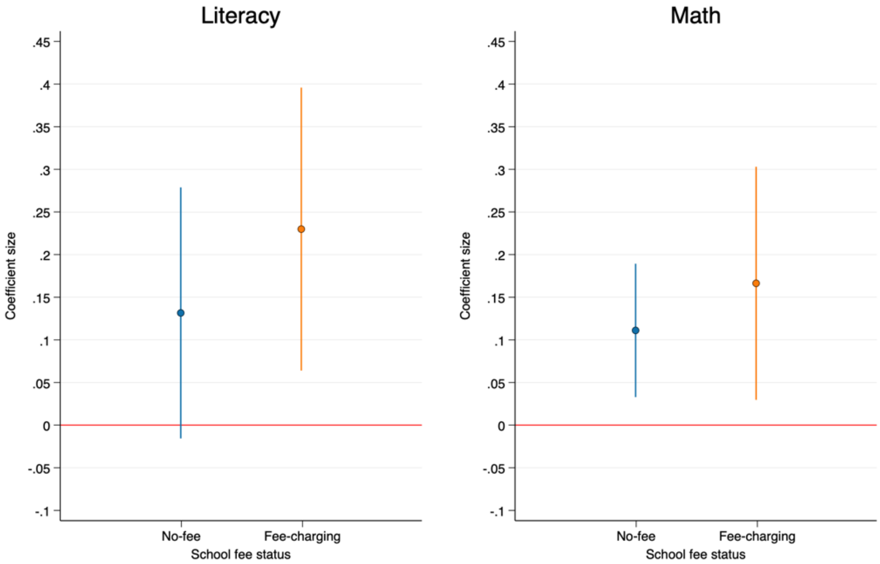


Notes: This figure shows the size of the coefficients on school-entry socio-emotional skills by subject and gender. The lines indicate the 90% confidence intervals. The regressions control for school-entry cognitive skills, age, SES, school fee status (no-fee or fee-charging), school language of instruction and school treatment status. School-entry socio-emotional skills predict later literacy and math achievement for boys but not girls. However, the confidence intervals are large, such that the differences are not statistically significant at the 90% level. Sample size: 195 boys; 205 girls.

5.3.2 School fee status

Yeager and Walton (2011) argue that socio-emotional skills interventions are not "magic" and that socio-emotional skills can support learning only when adequate learning opportunities exist within schools. Following this logic, one might expect school quality to mediate the association between school-entry socio-emotional skills and academic achievement. This is an important question from a policy perspective since efforts to foster early socio-emotional skills will not result in improved learning outcomes if these skills do not increase achievement in low-quality schools. The question can be investigated by using school fee status (no-fee or fee-charging) as a proxy for school quality and estimating Equation (1) separately for learners in no-fee and fee-charging schools. The estimated coefficients are shown in Figure 5. The results reveal important differences in the effect school-entry socio-emotional skills have on later academic achievement across school fee statuses. A 1-SD increase in school-entry socio-emotional skills is associated with a 0.23-SD increase in Grade 2 literacy scores for learners in fee-charging schools, whereas it is estimated to have a smaller and statistically insignificant 0.13-SD effect for learners in no-fee schools. This suggests that the academic returns on socio-emotional skills—in terms of literacy achievement at least—may be contingent on school-level factors such as teaching quality or classroom environment, which tend to differ systematically by school fee status. For mathematics, however, school-entry socio-emotional skills are significantly associated with academic achievement in both no-fee and fee-charging schools. This more consistent relationship may reflect the nature of early mathematics learning, which may be more directly influenced by behavioural and attentional capacity.

Figure 5: Conditional associations between school-entry socio-emotional skills and later academic achievement, by subject and school fee status



Notes: This figure shows the size of the coefficients on school-entry socio-emotional skills by subject and school fee status. The lines indicate the 90% confidence intervals. The regressions control for school-entry cognitive skills, age, gender, SES, school language of instruction and school treatment status. The coefficient is larger for learners in fee-charging schools than those in no-fee schools for both later literacy achievement and later mathematics achievement. However, the confidence intervals are large, such that the differences are not statistically significant at the 90% level. Sample size: 285 no-fee schools; 115 fee-charging schools.

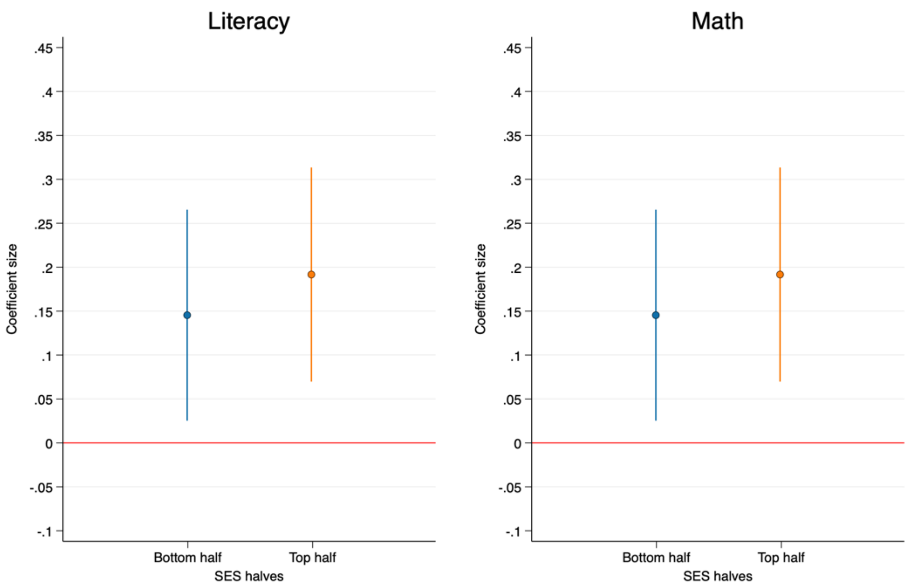
5.3.3 Socio-economic status

Another issue that is related to whether school quality affects the power of socio-emotional skills to predict later literacy and mathematics achievement is whether the predictive power of socio-emotional skills differs by learner SES. In line with the same logic laid out by Yeager and Walton (2011) that is mentioned in Section 5.3.2, learners from more disadvantaged backgrounds may experience smaller returns on their school-entry socio-emotional skills if they have fewer opportunities to leverage these skills at home to support learning. This possibility was investigated by splitting the sample into two groups based on SES[9] and estimating Equation (1) separately for each group. The resultant coefficients on school-entry socio-emotional skills are shown in Figure 6 and provide suggestive evidence in support of Yeager and Walton (2011)'s hypothesis, with larger coefficients estimated for learners in the top half of the SES distribution. It is noteworthy that the coefficients on

[9] SES was measured as the composite score on the asset index and obtained by applying PCA to the 13 assets included in the index.

socio-emotional skills are statistically significant for learners in both the lower and upper halves of the SES distribution. This indicates that school-entry socio-emotional skills are associated with later academic achievement across the socio-economic spectrum and may be particularly relevant for promoting educational equity given they are positively associated with achievement for learners from less-advantaged backgrounds.

Figure 6: Conditional associations between school-entry socio-emotional skills and later academic achievement, by subject and SES level

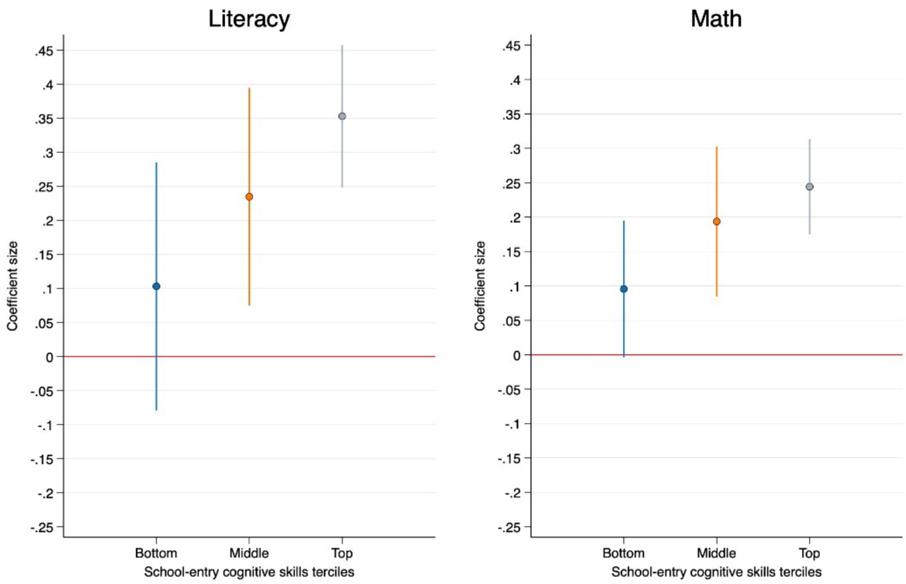


Notes: This figure shows the size of the coefficients on school-entry socio-emotional skills by subject and half of the SES distribution. The lines indicate the 90% confidence intervals. The regressions control for school-entry cognitive skills, age, gender, school fee status, school language of instruction and school treatment status. While the estimated coefficient is slightly larger for learners in the top half of the SES distribution than those in the bottom half for both subjects, school-entry socio-emotional skills nevertheless significantly predict later academic achievement for learners in the bottom half of the distribution. Sample size: 200 learners in the bottom half; 200 learners in the top half.

5.3.4 School-entry cognitive skills

Given the scaffolding nature of cognitive skills, it could be that learners who start school at an advantage in terms of their cognitive skill level may experience higher returns on their socio-emotional skills as they progress through the early grades and acquire foundational literacy and mathematics skills (Pollack et al., 2021). Estimating Equation (1) separately for learners with different cognitive skill levels at school entry provides evidence in support of this hypothesis, as shown in Figure 7. For learners in the bottom third of the distribution of baseline cognitive skills, not only is the estimated coefficient very small for both subjects, but the association is not statistically significant. In contrast, for learners in the middle and top thirds of the distribution, the coefficient is large and highly significant for both subjects. As with earlier findings, limited statistical power warrants a cautious interpretation of this finding. The results are nevertheless suggestive that the returns on socio-emotional skills are larger for learners who enter school with stronger cognitive foundations.

Figure 7: Conditional associations between school-entry socio-emotional skills and later academic achievement, by subject and school-entry cognitive skill level



Notes: This figure shows the size of the coefficients on school-entry socio-emotional skills by subject and third of the distribution of school-entry cognitive skills. The lines indicate the 90% confidence intervals. The regressions control for age, gender, SES, school fee status, school language of instruction and school treatment status. The coefficients are much larger for learners in the middle and top thirds of the distribution, which suggests that a certain level of cognitive skills is required for socio-emotional skills to be able to support the acquisition of academic skills in the early grades. Sample size: 134 learners in the bottom third, 133 learners in the middle third; 133 learners in the top third.

5.4 Theoretical mechanisms

There are several plausible mechanisms through which school-entry socio-emotional skills may influence later literacy and mathematics achievement. First, there is evidence that children with higher levels of socio-emotional skills form stronger relationships with their teachers (Birch and Ladd, 1998), which leads to them receiving more individualized support and encouragement (Graziano et al., 2007). Second, socio-emotional skills help children form supportive relationships with their peers, which also contributes to a more supportive learning environment (Wang et al., 2019). Lastly, children with stronger socio-emotional skills are better able to avoid distractions and persist with challenging tasks (Sorrenti et al., 2025). While these mechanisms provide valuable theoretical explanations for the relationship observed between early socio-emotional skills and later academic achievement, the data used do not make it possible to directly test these pathways. Future research with richer measures of classroom behaviour and teacher–student interactions would be needed to empirically validate these mechanisms.

VI.Conclusion

This study contributes to the growing body of literature on the role socio-emotional skills play in early learning by investigating their association with academic achievement in a longitudinal sample of South African learners. The results show that learners' socio-emotional skills at school entry are positively associated with their Grade 2 literacy and mathematics scores, even after controlling for school-entry cognitive skills and a number of other factors. The results of a sensitivity analysis suggest that the association between school-entry socio-emotional skills and later mathematics achievement is robust to omitted variable bias under the assumption of equal selection on observables and unobservables. However, the corresponding association with later literacy achievement is more sensitive to potential bias from unobserved confounders and should be interpreted with greater caution.

The results of a heterogeneity analysis reveal important differences in the strength of the association between school-entry socio-emotional skills and later academic achievement across subgroups. The results suggest that socio-emotional skills are more predictive of later achievement for boys than girls, particularly when it comes to literacy, and for learners in fee-charging schools than those in no-fee schools. Socio-emotional skills are also more strongly associated with later achievement among learners from higher socio-economic backgrounds and those with stronger cognitive skills at school entry. These patterns indicate that the extent to which socio-emotional skills support academic development may depend on the learning opportunities available to children both at home and at school. Nonetheless, school-entry socio-emotional skills are significantly associated with later academic achievement across the socio-economic spectrum, which underscores their relevance for learners from diverse backgrounds.

These findings have implications for both policy and research. They provide empirical motivation for integrating SEL into interventions aimed at strengthening foundational literacy and numeracy as has been done in other contexts (Jones, Brown and Aber, 2011; McCormick et al., 2021). Enhancing the socio-emotional climate of classrooms may also generate positive peer spillovers (McCormick et al., 2015) and support broader efforts to improve school quality.

Future research should examine the generalizability of these findings to larger and more diverse samples, ideally using experimental or quasi-experimental designs to better identify causal effects. While emerging evidence from high-income settings points to the long-term importance of socio-emotional skills for life outcomes, robust causal evidence from LMICs remains limited. A deeper understanding of how these skills interact with unequal educational opportunities in childhood and adolescence is essential for informing policies that promote more equitable learning and life trajectories.

Reference

Almlund, Mathilde, Angela Lee Duckworth, James Heckman and Tim Kautz. 2011. Personality psychology and economics. In *Handbook of the Economics of Education, Vol. 4*, ed. Eric A. Hanushek, Stephen J. Machin and Ludger Woessmann. Elsevier. pp. 1–158.

Altonji, Joseph G., Todd E. Elder and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113:151–184.

Ardington, Cally. 2022. Funda Wande Western Cape Evaluation: Baseline Report. Technical report. SALDRU, University of Cape Town.

Ardington, Cally, Gabrielle Wills, Elizabeth Pretorius, Nicola Deghaye, Nompumelelo Mohohlwane, Alicia Menendez, Nangamso Mtsatse and Servaas Van der Berg. 2020. Benchmarking early grade reading skills in Nguni languages. Technical report. Resep, Stellenbosch University.

Ardington, Cally, Nompumelelo Mohohlwane and Zelda Barends. 2022. Benchmarks Report: Afrikaans Early Grade Reading. Technical report. Department of Basic Education.

Barbarin, Oscar A. and Linda Richter. 2001. "Economic status, community danger and psychological problems among South African children." *Childhood* 8:115–133.

Birch, Sondra H. and Gary W. Ladd. 1998. "Children's Interpersonal Behaviors and the Teacher-Child Relationship." *Developmental Psychology* 34:934–946.

Böhmer, Bianca and Gabrielle Wills. 2023. COVID-19 and inequality in reading outcomes in South Africa: PIRLS 2016 and 2021. Working paper. Research on Socio-Economic Policy.

California Department of Education. 2008. California Preschool Learning Foundations, Volume 1.

California Department of Education. 2010. California Preschool Learning Foundations, Volume 2.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach and Danny Yagan. 2011. "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR." *The Quarterly Journal of Economics* 126(4):1593–1660.

Child Trends. 2014. Measuring elementary school students' social and emotional skills: Providing educators with tools to measure and monitor social and emotional skills that lead to academic success.

Cunha, Flavio, James J. Heckman, Lance Lochner and Dimitriy V. Masterov. 2006. "Interpreting the evidence on life cycle skill formation." *Handbook of the Economics of Education* 1:697–812.

Danner, Daniel, Clemens M. Lechner and Marion Spengler. 2021. "Editorial: Do we need socio-emotional skills?" *Frontiers in Psychology* 12:1155–1192.

Danon, Alice, Jishnu Das, Andreas de Barros and Deon Filmer. 2024. "Cognitive and socioemotional skills in low-income countries: Measurement and associations with schooling and earnings." *Journal of Development Economics* 168.

Dawes, Andrew, Linda Biersteker, Elizabeth Girdwood, Matthew Snelling and Colin Tedoux. 2016. Early Learning Outcomes Measure (ELOM) Technical Manual. Technical report.

Duncan, Greg J., Chantelle J. Dowsett, Amy Claessens, Katherine Magnuson, Aletha C. Huston, Pamela Klebanov, Linda S. Pagani, Leon Feinstein, Mimi Engel, Jeanne Brooks-Gunn, Holly Sexton, Kathryn Duckworth and Crista Japel. 2007. "School Readiness and Later Achievement." *Developmental Psychology* 43:1428–1446.

Durlak, Joseph A., Roger P. Weissberg, Allison B. Dymnicki, Rebecca D. Taylor and Kriston B. Schellinger. 2011. "The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions." *Child Development* 82(1):405–432.

Díaz, Mateo-Berganza, María Mercedes, Emma Näslund-Hadley, Margarita Cabra and Laura F. Vélez Medina. 2023. "Socioemotional Learning in Early Childhood Education: Experimental Evidence from the Think Equal Program's Implementation in Colombia." *IDB Working Paper*.

García, Jorge L., James J. Heckman, Duncan E. Leaf and María J. Prados. 2020. "Quantifying the life-cycle benefits of an influential early-childhood program." *Journal of Political Economy* 128(7):2502–2541.

Giese, S., A. Dawes, C. Tredoux, F. Mattes, G. Bridgman, S. Van der Berg, J. Schenk and J. Kotzé. 2022. Thrive by Five Index Report Revised August 2022. Report. Innovation Edge.

Graziano, Paulo A., Rachael D. Reavis, Susan P. Keane and Susan D. Calkins. 2007. "The role of emotion regulation in children's early academic success." *Journal of School Psychology* 45:3–19.

Heckman, James J. and Tim Kautz. 2012. "Hard evidence on soft skills." *Labour Economics* 19(4):451–464.

Heckman, James J. and Yona Rubinstein. 2001. "The Importance of Noncognitive Skills: Evidence from the GED Testing Program." *American Economic Review* 19:145–149.

Hofmeyr, Heleen, Cally Ardington and Nic Spaull. 2022. Roots and Shoots Baseline Report. Technical report.

Jones, Stephanie M., Joshua L. Brown and Lawrence Aber. 2011. "Two-Year Impacts of a Universal School-Based Social-Emotional and Literacy Intervention: An Experiment in Translational Developmental Research." *Child Development* 82:533–554.

Li-Grining, Christine P., Elizabeth Votruba-Drzal, Carolina Maldonado-Carreño and Kelly Haas. 2010. "Children's Early Approaches to Learning and Academic Trajectories Through Fifth Grade." *Developmental Psychology* 46:1062–1077.

Lopez Garcia, Italo, Jill Luoto, Frances Aboud and Lia Fernald. 2023. "Group Meetings and Boosters to Sustain Early Impacts on Child Development: Experimental Evidence from Kenya." *IZA Discussion Papers*.

McCormick, Meghan P., Elise Cappella, Erin E. O'Connor and Sandee McClowry. 2015. "Social-Emotional Learning and Academic Achievement: Using Causal Methods to Explore Classroom-Level Mechanisms." *AERA Open* 1.

McCormick, Meghan P., Robin Neuhaus, Erin E. O'Connor, Hope I. White, E. Parham Horn, Samantha Harding, Elise Cappella and Sandee McClowry. 2021. "Long-Term Effects of Social-Emotional Learning on Academic Skills: Evidence from a Randomized Trial of INSIGHTS." *Journal of Research on School Effectiveness* 14:1–27.

Moses, Eldridge and Servaas Van der Berg. 2023. "Evaluating the demand, supply and impact of early childhood development programmes in South Africa." *Development Southern Africa* 40:1153–1173.

Oster, Emily. 2019. "Unobservable selection and coefficient stability: Theory and evidence." *Journal of Business & Economic Statistics* 37:187–204.

Pollack, Courtney, Dayna Wilmot, Tracy M. Centanni, Kelly Halverson, Isabelle Frosch, Anila M. D'Mello, Rachel R. Romeo, Andrea Imhof, Jimmy Capella, Karolina Wade, Noor Z. Al Dahhan, John D. E. Gabrieli and Joanna A. Christodoulou. 2021. "Anxiety, Motivation, and Competence in Mathematics and Reading for Children With and Without Learning Difficulties." *Frontiers in Psychology* 12.

Pretorius, Elizabeth, Kathleen Rastle and Nangamso Mtsatse. 2022. A Curriculum Review of South African CAPS Grades R to 3. In *Early Grade Reading in South Africa*, ed. Nic Spaull and Elizabeth Pretorius. Cape Town: Oxford University Press Southern Africa. Chapter 7, pp. 143–163.

Pritchett, Lant. 2013. *The rebirth of education: Schooling ain't learning*. Centre for Global Development.

Richter, Linda and M.-L. Samuels. 2018. "The South African universal preschool year: A case study of policy development and implementation." *Child: Care, Health and Development* 44(1):12–18.

Roberts, Brent W., Nathan R. Kuncel, Rebecca Shiner, Avshalom Caspi and Lewis R. Goldberg. 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* 2:313–345.

Rodriguez-Segura, Daniel and Savannah Tierney. 2024. The Prevalence and Policy Implications of Between-School Heterogeneity in Learning Outcomes: Evidence from Six Public Education Systems. EdWorkingPaper 24-940. Annenberg Institute at Brown University.

Samuels, Alecia M., Wiedaad Slemming and Sadna Balton. 2012. "Early Childhood Intervention in South Africa in Relation to the Developmental Systems Model." *Infants and Young Children* 25:334–345.

Sorrenti, Guiseppe, Ulf Zölitz, Denis Ribeaud and Manuel Eisner. 2025. "The causal impact of socio-emotional skills training on educational success." *Review of Economic Studies* 00:1–47.

Spaull, Nic and Nwabisa Makaluza. 2019. "Girls do better: The pro-female gender gap in learning outcomes in South Africa 1995–2018." *Agenda* pp. 207–224.

Spaull, Nic and Stephen Taylor. 2022. Impact or scale? The trade-offs of early grade reading and mathematics interventions in South Africa. In *Early grade reading and mathematics interventions in South Africa*, ed. Nic Spaull and Stephen Taylor. Oxford University Press. pp. 1–23.

Todd, Petra E. and Kenneth I. Wolpin. 2003. "On the specification and estimation of the production function for cognitive achievement." *The Economic Journal* 113:F3–F33.

Venkat, Hamsa and Nicky Roberts. 2022. *Early Grade Mathematics in South Africa*. Oxford University Press Southern Africa.

Wang, Yehui, Zhaoxi Yang, Yingbin Zhang, Faming Wang, Tour Liu and Tao Xin. 2019. "The effect of social-emotional competency on child development in Western China." *Frontiers in Psychology* 10.

Wolf, Sharon and Dana Charles McCoy. 2019. "The role of executive function and social-emotional skills in the development of literacy and numeracy during preschool: a cross-lagged longitudinal study." *Developmental Science*.

World Bank. 2018. *World Development Report 2018: Learning to realize education's promise*. World Bank Group.

Yamauchi, Futoshi. 2011. "School quality, clustering and government subsidy in post-apartheid South Africa." *Economics of Education Review* 30:146–156.

Yeager, David S. and Gregory M. Walton. 2011. "Social-psychological interventions in education: They're not magic." *Review of Educational Research* 81:267–301.

Zehner, Tracy M., Tanya M. Paes, Brianna L. Devlin, Elyssa A. Geer, Germán Posada, Robert J. Duncan, David J. Purpura and Sara A. Schmitt. 2024. "Social–Emotional Competence as a Predictor of Early Numeracy Skills." *Early Education and Development* 35:1695–1711.

VII. Appendix

Sampling details

Afrikaans-medium schools

The 50 Afrikaans-medium schools included in the Roots and Shoots study were drawn from the sample used for the Funda Wande impact evaluation, which was conducted in collaboration with the Western Cape Education Department (WCED). The WCED selected 50 treatment schools to evaluate in accordance with criteria such as the presence of effective school management and the need to improve performance (Ardington, 2022). The treatment schools were matched with comparable control schools using administrative data on systemic test scores, fee status, and geographic location.

Of the 100 schools included in the Funda Wande evaluation, 50 schools (25 treatment and 25 control) located in the five education districts closest to Cape Town—Metro Central, Metro East, Metro North, Metro South and Cape Winelands—were selected for participation in Roots and Shoots. Selection was restricted to these districts to minimise fieldwork costs. Schools were stratified by district (with Metro Central, Metro East and Metro South combined into one stratum due to their low number of Afrikaans-medium schools) and by school fee status (no-fee vs. fee-charging) for a total of six strata. Four treatment schools were randomly selected from each stratum, and one additional school was selected from the Cape Winelands district to obtain a total of 25 treatment schools. These schools were then matched with 25 control schools based on systemic test performance to produce the final sample of 50 Afrikaans-medium schools. The Funda Wande evaluation report shows that the matching process resulted in treatment and control schools that were highly similar on a range of baseline characteristics at the school, teacher, and learner levels, with no statistically significant differences except for a few teacher-level variables (Ardington, 2022).

isiXhosa-medium schools

The 25 isiXhosa-medium schools were selected from the WCED's list of public schools located in the Metro East, Metro Central, Metro North, and Metro South districts. All the isiXhosa-medium schools in these districts are no-fee schools, so stratification was based on district and academic performance. The schools in each district were divided into performance quintiles on the basis of systemic test results. In Metro Central, Metro North, and Metro South, one school was selected from each quintile. In Metro East, which has a larger number of schools, two schools were selected per performance quintile. This resulted in a total sample of 25 isiXhosa-medium schools that reflect a range of performance levels.

Learner-level sampling

One Grade R class in each school was randomly selected to participate in the Roots and Shoots study. Eight children per class were randomly selected for assessment, for a total planned sample of 600 learners. Due to caregiver withdrawal of consent (N = 11) and unsuitable testing conditions (N = 33), the final Wave 1 sample consisted of 556 learners.

Items making up the ELOM Social and Emotional Functioning scale

Social relations with peers and adults

- Does this child work well with peers (can wait for their turn/manage impulsivity)?
- Does the child resolve problems with peers without becoming aggressive?
- Does the child cooperate with peers without prompting?
- Does the child seek out assistance or support from familiar adults?
- Does the child seek a familiar adult's ideas or explanations about events or experiences that are interesting to the child?
- Does the child take initiative in creating cooperative activities with a familiar adult?

Emotional readiness for school

- Is it easy to understand what the child is saying?
- Does the child express needs and feelings appropriately?
- Is the child independent, does the child like to do things without help?
- Does the child adjust well to changes in the classroom or home routine?
- Does the child approach new experiences confidently, without fear?
- Is the child a self-starter?

Table A1: OLS results for the effect school-entry socio-emotional skills have on later literacy and mathematics scores

	(1)	(2)	(3)	(4)	(5)	(6)
	Literacy	Math	Literacy	Math	Literacy	Math
Socio-emotional skills	0.334***	0.341***	0.141**	0.198***	0.248***	0.260***
	(0.065)	(0.060)	(0.062)	(0.055)	(0.068)	(0.062)
Cognitive skills			0.381***	0.406***	0.275***	0.373***
			(0.055)	(0.053)	(0.060)	(0.055)
Female			0.187**	-0.116	0.136	-0.142*
			(0.080)	(0.072)	(0.085)	(0.074)
Age			0.024	0.004	0.004	-0.001
			(0.017)	(0.015)	(0.018)	(0.017)
Social grant receipt			-0.009	-0.008	-0.023	-0.031
			(0.092)	(0.089)	(0.109)	(0.106)
Asset index			0.047	0.044	-0.038	-0.050
			(0.046)	(0.050)	(0.056)	(0.050)
Language of instruction			0.372***	-0.016		
			(0.133)	(0.141)		
Fee-charging school			0.378***	0.361**		
			(0.123)	(0.140)		
Treatment			0.184	-0.091		
			(0.132)	(0.133)		
Constant	0.004	0.004	-2.688**	-0.632	-0.327	0.166
	(0.071)	(0.072)	(1.116)	(0.908)	(1.159)	(1.095)
School FEs	No	No	No	No	Yes	Yes
N	400	400	400	400	400	400
R^2	0.113	0.118	0.318	0.335	0.244	0.305

Notes: School-entry socio-emotional skills and Grade 2 literacy and mathematics scores were standardized to have a mean of 0 and a standard deviation of 1. Standard errors are reported in parentheses. Asterisks indicate statistical significance at * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Models 1 and 2 estimate the effect in the absence of controls. Models 3 and 4 control for school-entry cognitive skills, gender, age, social grant receipt, asset index scores, school fee status, school language of instruction and school treatment status. Models 5 and 6 control for school-entry cognitive skills, gender, age, social grant receipt and asset index scores.

Table A2: Inverse probability weighting estimates of the association between school-entry socio-emotional skills and Grade 2 academic achievement

	(1) Literacy	(2) Math
Socio-emotional skills	0.115*	0.115***
	(0.066)	(0.039)
Cognitive skills	0.398***	0.283***
	(0.056)	(0.040)
Female	0.163**	-0.103*
	(0.079)	(0.055)
Age	0.031	-0.001
	(0.020)	(0.011)
Social grant receipt	-0.010	-0.027
	(0.097)	(0.060)
Asset index	0.017	0.025
	(0.050)	(0.037)
Language of instruction	0.466***	-0.005
	(0.145)	(0.112)
Fee-charging school	0.439***	0.272**
	(0.129)	(0.105)
Treatment	0.245*	-0.066
	(0.139)	(0.097)
Constant	-3.419**	0.331
	(1.316)	(0.650)
N	400	400
R^2	0.331	0.315

Notes: School-entry socio-emotional skills and Grade 2 literacy and mathematics scores were standardized to have a mean of 0 and a standard deviation of 1. Standard errors are reported in parentheses. Asterisks indicate statistical significance at * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$