EDUCATION
WORKING PAPER
No. 9 | JANUARY 2025

What's at Play? Unpacking the Relationship between Teaching and Learning

Brian Stacy, Maryam Akmal, Halsey Rogers,
Sergio Venegas Marin, Hersheena Rajaram,
and Viyaleta Farysheuskaya

Education Operations
SUPPORT HUB

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

1818 H Street NW
Washington DC 20433
Telephone: 202-473-1000
Email: AskEd@worldbank.org
Internet: www.worldbank.org/en/topic/education

Cover design: Marianne Siblini

What's at Play? Unpacking the Relationship between Teaching and Learning

Brian Stacy, Maryam Akmal, Halsey Rogers, Sergio Venegas Marin, Hersheena Rajaram, and Viyaleta Farysheuskaya[1]

Abstract

Using unique nationally representative school and system survey data from 13 education systems in low- and middle-income countries collected through the World Bank's Global Education Policy Dashboard (GEPD), we examine how the pedagogical practices, including practices to foster student engagement, and subject content knowledge of primary-school teachers correlate with their students' learning outcomes. We find that student performance on literacy (and to a lesser extent math) assessments is correlated with receiving instruction from teachers with better measured pedagogical skills. While the better-pedagogy effect is modest for the full sample, it is statistically robust and quite substantial for the upper-middle-income countries in the sample. Based on a sub-sample of those education systems, we also find that the use of learning strategies that support greater student engagement appears to be highly predictive of student learning outcomes in literacy. Better pedagogical practices are correlated with teachers' exposure to more practical, school-based pedagogical support—for example through induction or mentoring and feedback on lesson plans—and also with better teacher evaluation at the school level. The findings confirm the important role of interventions providing direct pedagogical support and feedback to teachers through training, instructional leadership, and evaluation, and they highlight the potential for interventions to foster student engagement and improve learning outcomes.

JEL: I21, J24, O15, C83

Keywords: Teacher pedagogy, student engagement, learning outcomes, education quality, human capital, teacher evaluation

[1] The order of author names was randomly assigned using the American Economic Association's author randomization tool. We would like to thank LEGO Foundation for funding for this paper. Comments from Ali Hasan Ansari, Dani Clark, Halil Dundar, Laura Gregory, Amer Hasan, Huma Kidwai, Ezequiel Molina, Lauri Pynnonen, Harshil Kumar Sahai, and participants in a seminar organized by LEGO Foundation and RTI are gratefully acknowledged. Thanks to Mohammed Reda Amin Eldesouky and Ahmed Raza for research assistance.

Contents

The Global Education Policy Dashboard: A Tool to Understand the Drivers of Learning Poverty

Over the past few decades, low- and middle-income countries have made significant strides in improving school enrollment rates (Ritchie et al. 2023). However, persistently poor learning outcomes cast a shadow over these advancements. Even before the COVID-19 pandemic, the learning poverty rate in 2019 was estimated at 57% in low- and middle-income countries, and as high as 86% in Sub-Saharan Africa (World Bank 2022). This means that a majority of children in these countries cannot read and understand a simple text by age 10.

The learning crisis is reflected in results from assessments conducted in different regions of the developing world. For instance, the PASEC (Programme d'Analyse des Systèmes Éducatifs de la CONFEMEN) 2019 data for 14 francophone countries in Sub-Saharan Africa reveals that only 48% and 38% of students in reading and mathematics, respectively, reach a sufficient competency level by the end of primary school (Confemen 2020). Similarly, the Annual Status of Education Report (ASER) reports that in India, only one in four Grade 3 students can perform tasks, such as basic subtraction, that are part of the second-grade curriculum (ASER 2023). Similarly, Early Grade Reading Assessment (EGRA) results from across various countries and languages show children struggling with basic reading sub-skills, such as letter name, letter sound, and word identification (Early Grade Reading Barometer 2024). Using literacy data from DHS (Demographic and Health Surveys) and MICS (Multiple Indicator Cluster Surveys), Le Nestour, Moscoviz & Sandefur (2022) find that literacy, conditional on completing five years of schooling, stagnated across the developing world over half a century, including absolute declines in both South Asia and Sub-Saharan Africa.

A critical question is how best to improve teaching so that children acquire the literacy and numeracy skills they need to become successful members of their communities. Based on lessons from countries that have achieved improvements in foundational literacy at scale, the World Bank's Literacy Policy Package highlights the critical ingredients to ensure that instruction in the classroom translates into learning gains for students. One key element is the role of effective and well-supported teachers, who have clear curricula and pedagogical guidance (for example, through effective initial teacher education and induction, teacher guides, and lesson plans or other resources), sufficient instructional time, and continuous support through coaching or other continuous professional development (Crawford et al. 2021). Similarly, an analysis of eight effective large-scale donor-led education programs in low- and middle-income countries highlights that the key elements influencing program success were often related to effective support for teachers, for example through practical training, teachers' guides, and follow-up coaching using structured tools, among others (Stern et al. 2021).

The World Bank's Global Education Policy Dashboard (GEPD) is uniquely positioned to answer questions about the relationship between teaching and learning because it deploys multiple instruments in one survey (for example, student assessments, teacher assessments, and the Teach Primary classroom observation tool, among others), covering the same students/teachers to paint a holistic picture of the correlation between teacher practices and student learning. Given that classroom observation videos are collected as part of the GEPD implementation, tools can also be applied ex-post. This was the case for this paper, where the recently developed PLAY tool (Jukes et al. 2022), now known as *Engage*[2], was also used to code classroom observation videos in order to gain insights into the prevalence of teaching practices to foster student engagement with learning, including through play-based and child-centered

[2] The name PLAY is used throughout this paper as it corresponds to the name used in publicly available materials at the time of writing. The authors acknowledge that the tool may be renamed to *Engage* in the future.

approaches, in classrooms around the world. Below we highlight the key components for effective teaching as defined in the GEPD framework and encompassing instruments.

The GEPD framework on essential ingredients for effective teaching

To enable learning, teachers must:	**Captured through GEPD survey questions on:**
Be motivated and be present	Teacher intrinsic motivation and presence
Know the content they are teaching	Teacher subject content knowledge assessment
Teach it well	Teach Primary and PLAY scores
Be supported with training	Teacher training, follow on support, and evaluation
Be attracted to the teaching profession	Attraction, selection, and deployment
Have incentives to improve performance	Monitoring, accountability, and intrinsic motivation

Leveraging GEPD data from 13 education systems, the paper first establishes what we know about student learning outcomes, teacher pedagogical practices, subject content knowledge (hereafter referred to as content knowledge), and how well-supported teachers are. Building on these findings, the paper addresses the following key questions:

- How do teacher pedagogical practices and content knowledge correlate with student learning?
- How do teaching practices focused on fostering student engagement with learning, including through play, correlate with student learning?
- What factors predict better teaching quality?

The paper first provides an overview of existing literature on the mechanisms driving the relationship between teaching and learning. In the subsequent sections we describe our data, sample, and methodology. Further sections discuss the results. Finally, we conclude with a brief discussion of the implications for policy and further research.

Existing Literature Underscores the Role of Teachers in Improving Student Learning Outcomes

A large body of evidence highlights teaching quality as a key determinant of student learning. Existing research shows a relatively weak link between observable teacher characteristics (such as qualifications and experience) and teacher value-added, not because those characteristics are not important but because their quality is often too low to impact student learning (Evans & Beteille 2021). However, the quality of teacher-child interactions has been shown to explain a significant share of student learning (Dobbie & Fryer 2013; Muijs et al. 2014). Going from low-performing teaching to high-performing teaching (in terms of overall value-added) can substantially improve student learning, with estimated impacts ranging from 0.2 standard deviations in Ecuador to more than 0.9 standard deviations in India, which is equivalent to multiple years of schooling (Evans & Beteille 2021).

In fact, some of the most effective interventions to improve learning in low- and middle-income countries (LICs and MICs), such as structured pedagogy, work through teachers (Snilstveit et al. 2015). Many of the most effective reading programs in diverse contexts such as India, Pakistan, Kenya, Ghana, Nigeria, and Senegal encompass elements related to structured pedagogy that are focused on supporting teachers deliver effective instruction inside the classroom (Stern et al. 2021). The Global Education Evidence Advisory Panel (GEEAP) also highlights interventions supporting teachers with structured pedagogy as cost-effective and evidence-backed (Akyeampong et al. 2023).

The research on the link between teaching and student learning covers various different dimensions of teaching quality, ranging from teacher content knowledge to specific pedagogical practices to teacher presence and instructional time. To understand the impact of teacher effectiveness on student learning outcomes, it is important to disentangle the mechanisms driving student-teacher interactions that most affect student learning.

In many low- and middle-income countries, teachers often lack mastery of the content they are teaching. Bold et al. (2017) find that teachers' subject knowledge in primary schools in Africa is strikingly low. Only about one in ten Grade 4 teachers master their students' language curriculum, and about a quarter of the teachers fail at simple tasks, such as subtracting two-digit numbers or choosing the correct pronoun or conjunction to complete a sentence.

The content knowledge that teachers possess matters for student learning. Evidence shows that when students in low- and middle-income countries are in classes led by teachers with higher content knowledge, they learn more (Metzler & Woessmann 2012; Bietenbeck et al. 2017). Bold et al. (2019) show that deficits in teachers' content knowledge account for a substantial portion—30%—of the learning shortfall relative to the curriculum. Moreover, disparities in teachers' content knowledge explain roughly a fifth of the learning gap between the lowest and highest-performing countries. Bold et al (2019) find that a 1-standard-deviation increase in teacher content knowledge increases student learning by 0.06 standard deviations in the short run, meaning that moving a student from a teacher at the 5th percentile to the 95th percentile in terms of teacher content knowledge increases student learning by 0.2 standard deviations (SDs) in one year. Further reinforcing this, Filmer et al. (2020) find that enhancing subject content knowledge among teachers, coupled with increased classroom instructional time, correlates with improved student test scores. Going from low-performing to high-performing teacher in terms of assessed content knowledge is associated with a 0.17-SD increase in student test scores.

Achieving content mastery is only one part of the job. Teachers also need to be able to employ effective instructional practices to teach the content. There are fewer studies of teacher's pedagogical skills, and the little data that exists is not encouraging. The World Bank's Service Delivery Indicator (SDI) survey reveals that a considerable number of teachers, even in better-performing countries in Sub-Saharan Africa like Kenya and Tanzania, struggle with effective pedagogy, with only a third answering pedagogical questions correctly. Despite having adequate content knowledge, many teachers—such as those in Bihar, India, and Afghanistan —do poorly in explaining concepts, contributing to low levels of student learning (Beteille & Evans 2021).

Bold et al. (2017) highlight deficiencies in teachers' ability to evaluate student progress and employ effective teaching practices such as regular comprehension checks and feedback. Their analysis, leveraging student-teacher data across countries, emphasizes the substantial positive effects of both teacher content and pedagogical knowledge on student achievement. In their data, a 1-SD change in teacher content knowledge, pedagogical knowledge, and skills is associated with roughly a quarter of a standard deviation increase in student test scores. In Pakistan, Molina et al. 2020 find that a 1-SD increase in Teach scores in associated with a 0.068- to 0.124-SD increase in student test scores. In a Ugandan-

based randomized evaluation, Buhl-Wiggers (2017) found that a 1-SD increase in teaching quality is correlated with a notable 0.14-SD enhancement in student performance on year-end reading tests.

However, the effectiveness of teacher's content knowledge and pedagogical skills still hinges on their presence in the classroom. Research in African primary schools reveals that students receive only about half of the scheduled teaching time due to high levels of teacher absence (Bold et al. 2017). This absence significantly impedes student learning and results in considerable instructional time loss throughout the school year (Chaudhury et al. 2006; Lavy 2015). Experimental interventions in India showcase the potential impact of reducing teacher absence, with a 21-percentage-point decrease in absence leading to a 0.17 SD increase in student learning (Duflo et al. 2012).

What Do GEPD Results Tell Us About Students and Teachers?

The Global Education Policy Dashboard (GEPD) uses three data collection instruments to report on 39 indicators that, operationalizing the World Development Report 2018 framework, provide a snapshot of how the education system is working. The three instruments—School Survey, a Policy Survey, and a Survey of Public Officials—were built based on the best existing tools, streamlined to reduce the time and cost required to carry them out:

- GEPD School Survey – The School Survey builds on Service Delivery Indicators (SDI), Teach Primary, Global Early Child Development Database (GECDD), and the Development World Management Survey (DWMS), among other tools. This survey is carried out in a representative random national sample of schools and takes approximately four hours to complete per school. It includes eight modules: School Information, Roster, Principal Questionnaire, Teacher Questionnaire, Classroom Observation, and Assessments for Teachers and for Grade 1 and Grade 4 Students.
- GEPD Policy Survey – This survey builds on the SABER policy-data collection tools that the World Bank first introduced in 2010. It involves a legislative review conducted by a senior consultant in each country, drawing on their knowledge to identify the country's choices in policy areas highlighted in the dashboard.
- GEPD Survey of Public Officials – The public-officials survey, developed in collaboration with World Bank's Bureaucracy Lab, uses 45-minute interviews with education officials at the central and subnational levels. The sample of surveyed officials is representative at the targeted levels.

This paper largely focuses on the GEPD School Survey for information on schools, teachers (characteristics, presence, content knowledge, and pedagogical skills), students (characteristics and student knowledge), as well as the de facto policy environment in which school actors operate.

The aim of the GEPD School Survey is to produce nationally representative estimates with enough precision to allow detection of changes over time at a minimum power of 80% and at a 0.05 significance level. The sample size also allows for disaggregation by rural/urban location and by gender on relevant indicators. For the GEPD School Survey, a two-stage random sample design is used. In the first stage, a representative random sample of 200-300 schools is drawn per country. We use Probability-Proportional-to-Size (PPS) sampling based on the number of students enrolled at each school, after stratifying the sampling frame based on urban/rural classification and the region in which schools are located. In the second stage, a random sample of teachers and students within each school is selected to answer questions from the survey modules; this sampling is done by enumerators in the field. In every school, the principal

is interviewed, and the infrastructure is assessed by the enumerators. Ten teachers per school are sampled randomly for attendance checks, and five teachers are interviewed and given a teacher content knowledge assessment, where teachers are asked to mark and correct a hypothetical student's exam in math or language, depending on the subject they teach. Of these 5, one Grade 4 teacher is selected for a classroom observation.[3] The classroom observation entails the filming of a 45-minute lesson, which is later coded using the Teach Primary classroom observation tool. The students in the same Grade 4 class are also assessed for student learning on literacy and mathematics. Lastly, three Grade 1 students are randomly selected for a 1-on-1 assessment on literacy, numeracy, executive function, and socioemotional skills.

In this paper, we use data from 13 education systems: Chad (2023), Ethiopia (2020), Jordan (2019), Gabon (2023), Niger (2022), Pakistan's Punjab (2023), Khyber Pakhtunkhwa (KP, 2022), Balochistan (2023), and Sindh (2023) provinces and Islamabad Capital Territory (ICT, 2022), Peru (2019), Rwanda (2020), and Sierra Leone (2022). We further divide our sample of countries into income groups based on the latest World Bank classification. Low-income countries include Ethiopia, Niger, Sierra Leone, Rwanda, and Chad, lower-middle-income countries include Jordan and Pakistan, and upper-middle-income countries include Peru and Gabon. Across these 13 education systems, we sampled 2,885 schools employing 29,069 teachers; 2,686 teachers were selected for the classroom observation module and 2,144 teachers completed both the classroom observation module and the content knowledge assessment. In Grade 4, 50,392 students were tested for proficiency in literacy and numeracy.

To explore questions pertaining to practices that support student engagement with learning, including through play-based approaches, the same classroom observation videos coded with Teach in Peru, Ethiopia, and Sierra Leone were later coded with the classroom observation tool part of the broader Playful Learning Across the Years (PLAY) toolkit. This tool was jointly developed by LEGO Foundation, RTI International, and New York University (NYU) Global TIES for Children. The toolkit aims to measure how adults support children's self-sustaining engagement in learning, which leads to a broad range of learning outcomes. Before PLAY was created, no measurement tools had been available to assess children's engagement with learning in the classrooms (Jukes 2023). Given the tool's novel focus on fostering engagement, this study explores the link between such engagement and learning outcomes.

Table 1 shows the characteristics of schools in our sample, including the proportion of schools in rural areas, the number of students enrolled in schools, the average teacher absence rate (a teacher is considered absent if they were not in school on the day the enumerator visited), if a school had access to drinking water, a functioning toilet, internet access, and electricity in the classroom, and school monitoring. School monitoring measures the extent to which there is a monitoring system in place to ensure that the inputs that must be available at the schools are in fact available. The score ranges from 0 to 3, and includes 1 point if all input items (functioning blackboard, chalk, pens, pencils, textbooks, exercise books in Grade 4 classrooms, basic classroom furniture, and at least one computer in the schools) are being monitored, 1 point if all infrastructure items (functioning toilets, electricity, drinking water, and accessibility for people with disabilities) are being monitored, and 1 point if the community is involved in monitoring.

[3] Due to the significant time commitment required to complete both the content knowledge assessment and the classroom observation module, some teachers selected for the classroom observation module were not able to complete the content knowledge assessment (see Table A5 of Appendix A).

Table 1: School Characteristics

		Low-income					Lower-middle income						Upper-middle income	
	All	Ethiopia	Niger	Rwanda	Sierra Leone	Chad	Jordan	Pakistan ICT	Pakistan KP	Pakistan Balochistan	Pakistan Sindh	Pakistan Punjab	Peru	Gabon
Number of sampled schools	2885	299	270	200	302	262	250	100	200	200	200	200	206	196
Total number of teachers	29069	4633	2359	1584	2367	2071	4111	1863	1442	2044	1368	1788	1461	1978
% rural schools	59	15	93	85	39	57	22	89	94	59	90	91	19	1
	(1.5)	(2.4)	(.6)	(0)	(2.6)	(4)	(1.4)	(1.7)	(1)	(3.2)	(1.1)	(1.3)	(2.7)	(.3)
	[49]	[36]	[26]	[36]	[49]	[50]	[42]	[31]	[23]	[49]	[31]	[29]	[39]	[9]
Students enrolled (school level)	413	1180	238	888	272	336	659	367	163	192	135	282	479	174
	(18.1)	(107.2)	(11.1)	(35.2)	(11.5)	(22.2)	(56)	(35.9)	(12.6)	(25.2)	(7.6)	(15.8)	(93.9)	(26.5)
	[484]	[798]	[165]	[540]	[169]	[283]	[508]	[227]	[166]	[283]	[133]	[304]	[472]	[196]
School monitoring	3	3	2	3	3	2	3	4	3	2	2	3	3	3
	(0)	(.1)	(.1)	(0)	(.1)	(.1)	(0)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.2)
	[1]	[1]	[1]	[1]	[1]	[1]	[1]	[0]	[1]	[1]	[1]	[1]	[1]	[1]
Avg teacher absence rate (school level)	14	30	17	13	28	9	7	12	10	18	8	10	8	11
	(.8)	(5.3)	(2.1)	(.9)	(2)	(1.9)	(.9)	(2)	(1.7)	(2.8)	(2.4)	(1.4)	(1.5)	(3.3)
	[21]	[31]	[24]	[13]	[23]	[19]	[10]	[11]	[17]	[27]	[21]	[17]	[13]	[17]
Drinking water	70	63	54	74	64	34	89	93	71	39	50	96	91	92
	(1.3)	(4.6)	(4.9)	(3.2)	(3.9)	(4.9)	(2.7)	(4.2)	(5.6)	(4.2)	(4.5)	(1.5)	(2)	(4.4)
	[46]	[48]	[50]	[44]	[48]	[47]	[31]	[26]	[45]	[49]	[50]	[20]	[29]	[28]
Functioning toilet	34	34	15	57	32	14	69	53	28	7	3	45	68	21
	(1.4)	(6.2)	(3.5)	(2.7)	(3.5)	(2.7)	(4.3)	(7.4)	(4.3)	(1.8)	(1)	(4.3)	(5)	(7.1)
	[47]	[47]	[36]	[50]	[47]	[35]	[46]	[50]	[45]	[26]	[16]	[50]	[47]	[41]
Internet access	25	14	0	27	1	1	88	54	1	1	0	57	54	25
	(1.4)	(3.1)	(.2)	(2.9)	(.4)	(.9)	(1.9)	(8.8)	(.4)	(.4)	(.1)	(4)	(5.2)	(8.8)
	[42]	[30]	[3]	[43]	[8]	[10]	[32]	[48]	[7]	[10]	[3]	[49]	[48]	[42]
Class electricity	60	56	6	78	8	8	100	100	77	31	30	98	96	95
	(1.3)	(5.4)	(1.7)	(2.7)	(1.9)	(2.3)	(.5)	(.2)	(5.1)	(3.7)	(3.6)	(1.7)	(1.2)	(3.1)
	[49]	[50]	[24]	[42]	[28]	[28]	[7]	[4]	[42]	[46]	[46]	[16]	[20]	[21]

Table shows average with standard error in parentheses and standard deviation in brackets. Values in the overall column are the unweighted average of the country values. Values in country columns are weighted.

Table 2 describes teacher characteristics of the 2,686 teachers for whom we have information regarding their pedagogical skills. The average age of teachers in our sample ranges between 33 and 48 years old across countries, and the gender ratio of teachers varies substantially by country. While most teachers were men in Jordan, Peru, Ethiopia and Sierra Leone, the majority of teachers were female in Chad, Gabon, KP and Sindh. Overall, over 40% of teachers in the sample did not have a bachelor's or master's degree, highlighting the potential role of higher education in building teachers' content knowledge and pedagogical skills. Thirty-three percent of all teachers in our sample have obtained a bachelor's degree as their highest degree, and a further 23% have at least a master's degree. The highest level of education completed varies by country, with a higher proportion of teachers holding a master's degree in Pakistan. Refer to Table A1 in Annex A for teacher characteristics by country.

Table 2: Teacher Characteristics

Characteristic	Average	Min	Max
Number of teachers	206	100 (Pakistan ICT)	287 (Ethiopia)
Age	38.9	33.4 (Ethiopia)	47.7 (Peru)
Teaching experience (in years)	13.6	7.6 (Pakistan KP)	21.5 (Peru)
% male	53	12.4 (Chad)	86.6 (Jordan)
Highest level of education completed			
No education	0.1	0 (Ethiopia, Gabon, Jordan, Niger, Pakistan Balochistan, Pakistan Sindh, Pakistan ICT, Pakistan KP, Peru, Rwanda, Sierra Leone)	0.7 (Pakistan Punjab)
Less than primary	0	0 (Ethiopia, Gabon, Jordan, Pakistan Balochistan, Pakistan Punjab, Pakistan Sindh, Pakistan ICT, Pakistan KP, Peru, Rwanda, Sierra Leone, Chad)	0.1 (Niger)
Primary completed	0.3	0 (Jordan, Pakistan Balochistan, Pakistan Punjab, Pakistan Sindh, Pakistan ICT, Pakistan KP, Peru, Rwanda)	2.4 (Gabon)
Middle passed	4	0 (Jordan, Pakistan Balochistan, Pakistan Punjab, Pakistan Sindh, Pakistan ICT, Pakistan KP, Peru)	34.2 (Niger)
High school passed	22.2	0 (Jordan, Pakistan ICT)	90.7 (Rwanda)
Bachelor's degree	32.7	4.2 (Sierra Leone)	79 (Jordan)
Master's degree	22.3	0 (Ethiopia, Niger, Rwanda, Sierra Leone)	78.4 (Pakistan ICT)
Doctoral degree	0.7	0 (Gabon, Jordan, Niger, Pakistan Balochistan, Rwanda, Sierra Leone, Chad)	3.9 (Peru)
Teachers Training Institute (TTI) Graduate	0.6		7.1 (Ethiopia)
Teachers Training College (TTC) Graduate	5.2		62.5 (Ethiopia)
Other	11.9	0 (Pakistan Punjab, Pakistan Sindh, Pakistan ICT)	45.3 (Niger)
Subjects taught			
Language	19.9	0 (Gabon, Peru)	67.1 (Rwanda)
Mathematics	9.3	0 (Niger, Peru)	41.4 (Ethiopia)
Both/All	68.8	9.9 (Rwanda)	99.9 (Gabon)
Other	2	0 (Gabon, Niger, Rwanda, Sierra Leone)	16.4 (Jordan)

In Table 3, we present the student characteristics of the 50,392 Grade 4 students in our sample. The GEPD school surveys collect data on socio-economic aspects related to whether parents went to university, number of times a student ate in a day, number of pairs of shoes student owns, and the type of bathroom they have at home. This data is described in Table 3. However, since data on these components and on household size are not available for all countries (because they were not collected in the first two GEPD countries, Sierra Leone and Peru), to control for socio-economic status in later regressions, we instead use a school-level wealth index that draws on micro-estimates of relative poverty and wealth across countries using satellite imagery, mobile phone network data, and topographic maps, as well as nationally representative household survey data (Chi et al. 2021).[4] Refer to Table A2 in Annex A for student characteristics by country.

Across the 13 education systems, there are important differences. For instance, the average age of Grade 4 students varies significantly, from around 10 in Jordan to much higher in countries like Ethiopia (due to late enrollment or repetition). Similarly, there are substantial differences in household wealth, as proxied by aspects like the type of bathroom used at home and number of meals eaten in a day. For the sample overall, 7% of students report having no bathroom at home, and 42% of students report eating less than 3 times a day.

Table 3: Student Characteristics

Characteristic	Average	Min	Max
Number of students	3903	2350 (Pakistan Sindh)	5783 (Ethiopia)
Age	10.6	9.5 (Jordan)	11.6 (Ethiopia)
Household size	6.7	5.7 (Ethiopia)	7.6 (Niger)
Parents/caregivers went to university	46.6	29.9 (Pakistan Sindh)	83.4 (Gabon)
Number of times student ate			
Did not eat	.5	0 (Pakistan Balochistan, Pakistan Punjab)	3.3 (Ethiopia)
One to two times a day	41.9	19.2 (Ethiopia)	63.7 (Rwanda)
At least three times	57.7	36.3 (Rwanda)	77.5 (Ethiopia)
Number of pairs of shoes owned by a student			
Zero pairs of shoes	.5	0 (Pakistan Balochistan, Pakistan Punjab, Pakistan Sindh)	4 (Ethiopia)
One pair of shoes	21.3	4.4 (Gabon)	44.4 (Pakistan Sindh)
At least two pairs of shoes	78.2	55.6 (Pakistan Sindh)	95.6 (Gabon)
Type of bathroom at home			
Toilet flush	24	5.8 (Ethiopia)	71 (Jordan)
Pit	28.2	5.7 (Gabon)	75.3 (Rwanda)
Bucket	5.1	.1 (Jordan)	10.6 (Pakistan Sindh)
Communal bathroom	21.1	9.2 (Jordan)	36.5 (Pakistan Punjab)
No bathroom (outside)	7.3	.2 (Jordan)	26.9 (Niger)
No response	14.3	1.6 (Rwanda)	30.4 (Chad)

Information about household size, if parents went to university, number of times student ate per day, number of pairs of shoes they own, and the type of bathroom they have at home were not asked in Sierra Leone and Peru.

[4] For the GEPD observations that include data on SES, the wealth index has a substantial and highly significant correlation with the GEPD-based SES index (an index comprised of SES characteristics like number of shoes owned, type of bathroom in the home, parental education, and number of meals). The coefficients on the key variables of interest are also quite similar whether we use the SES index or the more complete wealth index to proxy for socioeconomic status.

Student Learning Outcomes Are Low Across the Board

The GEPD School Survey includes a Grade 4 student assessment.[5] Students in Grade 4 are assessed in literacy and mathematics, with each student receiving a score between 0 and 100. Student learning is thus defined by the score a student received in their assessment. The literacy assessment includes 17 questions on letter identification, word recognition, and reading comprehension, and the mathematics assessment includes 15 questions on number sense, arithmetic, word problems, and sequences. A student can receive partial credit for getting some questions right within a cluster of questions.[6] The literacy student knowledge score measures the average score a student received across the 6 questions[7] in the literacy assessment, and the mathematics student knowledge score measures the average score a student received across the 15 questions in the mathematics assessment. Student knowledge score (0-100) is the average of the student's literacy and mathematics knowledge scores. This paper focuses on the 50,392 students in Grade 4 who were tested in proficiency in literacy and numeracy. Table 4 presents the knowledge scores and individual items from the literacy and mathematics assessments.

Across education systems, students scored higher in literacy (65% of questions correct, on average) than in mathematics (40%). In the literacy assessment, students were more successful at linking pictures, letter recognition and word recognition. In the mathematics assessment, results varied more across the different questions. Across all education systems, students scored relatively higher on addition and subtraction questions and scored relatively lower on multiplication, division, and data interpretation questions. These patterns show that systems are succeeding in developing some of the earliest foundational literacy and numeracy skills of many children, but even those skills remain far from universal, and systems have less success in developing some of the higher-order skills such as reading comprehension. Of all education systems surveyed, students in Peru performed better across most items.

Table 4: Student Knowledge

Characteristic	Average	Min	Max
Number of students	3903	2350 (Pakistan Sindh)	5783 (Ethiopia)
Student knowledge	52.5	36 (Chad)	80.7 (Peru)
Literacy student knowledge	64.7	40.3 (Chad)	92.4 (Peru)
Mathematics student knowledge	40.3	28.5 (Rwanda)	69.1 (Peru)
Literacy skills			
Pupil can link pictures	81.2	41.1 (Niger)	99.4 (Peru)
Pupil can understand a story (1)	51.9	21.6 (Chad)	95 (Peru)
Pupil can understand a story (2)	52.5	24 (Pakistan KP)	90.2 (Peru)
Pupil can select a correct word	33.4	9 (Niger)	79.5 (Peru)
Pupil can recognize a letter	83.6	68.3 (Rwanda)	95.6 (Pakistan ICT)
Pupil can recognize a word	85.6	68.6 (Chad)	96.1 (Peru)
Mathematics skills			
Pupil can recognize numbers	88.9	80.4 (Chad)	98.3 (Peru)
Pupil can order numbers	42.1	15.4 (Chad)	81.3 (Peru)
Pupil can add single digits	76	62.5 (Pakistan Sindh)	94.1 (Peru)
Pupil can add double digits	61.3	46 (Rwanda)	90.1 (Peru)
Pupil can add triple digits	56.2	38.8 (Chad)	91 (Peru)
Pupil can subtract single digits	69.2	52.6 (Pakistan Sindh)	89.6 (Peru)

[5] Grade 4 Assessment is based on SDI Assessment, which was checked against the curriculum of 13 countries. Questions were added to better align to Global Proficiency Framework. For more information, refer to the GEPD Technical Guide.

[6] There are three clusters of questions in the literacy assessment (recognizing a word, recognizing a letter, and linking pictures), and one cluster of question in the math assessment where a student can receive partial credit (recognizing numbers).

[7] The literacy student knowledge assessment in Peru did not include one of the reading comprehension texts, which included 11 follow-up questions. Consequently, to compute the final literacy student knowledge score, we used only the 6 questions that were implemented across all sampled countries.

Pupil can subtract double digits	39.9	24.3 (Rwanda)	64.1 (Peru)
Pupil can multiply single digits	38.6	21.9 (Pakistan Sindh)	74 (Peru)
Pupil can multiply double digits	13.8	3 (Jordan)	45.4 (Peru)
Pupil can multiply triple digits	10	1.5 (Jordan)	39.2 (Peru)
Pupil can divide single digits	36.8	19.8 (Rwanda)	73 (Peru)
Pupil can divide double digits	22.7	10.3 (Rwanda)	54.7 (Peru)
Pupil understands division	12.9	2.4 (Rwanda)	34.8 (Peru)
Pupil can solve math story	14.5	2.3 (Rwanda)	45.5 (Peru)
Pupil can complete a sequence	21.6	2.6 (Rwanda)	61.3 (Peru)

Student knowledge scores is the average of literacy and mathematics knowledge scores. Literacy student knowledge score measures the average score each student received across the 24 questions in the assessment. Mathematics student knowledge score measures the average score each student received across 15 mathematics questions.

Teacher Pedagogical Practices Are Lacking and Content Knowledge is Low

Quality of teaching rests on a teacher's knowledge of the subject they teach and their ability to convey it in a way that students can understand. Through the GEPD, teaching quality is captured through the use of the classroom observation tool Teach Primary to measure the quality of pedagogical skills and teacher content knowledge assessment to assess teachers' familiarity with the content they teach. Teach Primary is an open-access classroom observation tool that was developed to measure the quality of teaching practices of primary school teachers in low-and-middle-income countries. Teach Primary is composed of three main areas: Classroom Culture, Instruction, and Socioemotional Skills. Each area can be further broken down to 3 or 4 elements, for a total of 10 elements. Each element is scored on a 1-to-5 scale and is further broken down into two to four corresponding practices, for a total of 27 practices, which are scored "low," "medium," or "high." A Teach Primary score (1-5) measures the average of the 27 practices and each subconstruct measures the average of the practices within. See Table 5 for a summary of the data collected across the 13 education systems. Table A4 in Annex A provides the same information disaggregated by education system.

Table 5: Teacher Pedagogical Skills

Characteristic	Average	Min	Max
Number of teachers	206.6	100 (Pakistan ICT)	287 (Ethiopia)
TEACH score	2.7	2.4 (Ethiopia)	3.3 (Gabon)
Supportive learning environment	3.3	2.9 (Sierra Leone)	4 (Peru)
Positive behavioral expectations	3	2.4 (Ethiopia)	3.6 (Gabon)
Lesson facilitation	3.1	2.5 (Pakistan Balochistan)	3.6 (Pakistan ICT)
Checks for understanding	2.9	2.5 (Chad)	3.9 (Gabon)
Feedback	2.5	1.6 (Rwanda)	4 (Gabon)
Critical thinking	2.6	2.1 (Rwanda)	3.7 (Gabon)
Autonomy	2.7	2.3 (Sierra Leone)	3.2 (Gabon)
Perseverance	2.3	1.7 (Chad)	2.9 (Pakistan ICT)
Social and collaborative skills	2	1.1 (Chad)	3 (Jordan)

Sample is restricted to teachers selected for the classroom observation module.

Across all systems, Teach scores tended to be higher for the classroom culture area, with teachers showcasing relatively good command of the practices and skills needed to foster supportive learning environments and outlining positive behavior expectations. Teachers exhibit supportive learning environment practices when their students feel supported in their learning and are encouraged to meet high academic and behavioral standards. Despite the higher average scores, there were disparities across the systems, with the gap between the maximum and minimum observed in those elements ranging from 0.9 to 2.4 points, on a 1-to-5 scale.

Beyond classroom culture, the education systems on average performed relatively well on some of the elements of quality of instruction (again with substantial disparities across education systems). Scores were relatively high for lesson facilitation, which occurs when teachers make lesson objectives clear, deliver content easily understandable by students, and connect content to material already familiar to the students. Education systems did not perform as well on other pedagogical practices related to instruction, with greater disparities across the different systems. In particular, education systems scored low on average on the provision of feedback, with the difference between the lowest and highest averages observed being the largest.

Similarly, pedagogical practices associated with the development of students' socioemotional skills appear to be relatively less common across all the surveyed systems, with the average score in all of the associated subconstructs being 2 on a 1-to-5 scale. These practices include aspects such as fostering student autonomy, promoting perseverance, and facilitating collaboration among the students. While social and collaborative skills refer to peer-to-peer interactions, collaboration, and cooperative learning, autonomy reflects students having a sense of control over their goals, thus being more engaged and motivated to learn. Perseverance captures students pushing through challenges by helping them understand that their potential can be developed, providing them with strategies for developing their potential, and reassuring them when setbacks happen.

Along with effective pedagogical skills, teachers need good mastery over the content they teach to be able to deliver learning. Language and mathematics teachers teaching Grade 4 in the current year (or previous year) were assessed in their respective subjects. The literacy assessment includes 20 questions on identifying a correct letter, cloze, grammar, and reading comprehension. The mathematics assessment includes 23 questions on arithmetic checks, geometry, and data interpretation. Up to 5 teachers were tested in each school. However, only 1 (at most) overlapped with the classroom observation. In some situations, due to lack of time on the teacher's side or issues of logistics within the school, the teacher who was observed was not able to participate in the content knowledge assessment. For the analysis pertaining to the relationship between teacher content knowledge and student learning, only observations where the assessed teacher overlaps with the observed teacher were included. As part of the test, teachers were asked to mark (or "grade") mock student tests in language and in mathematics. This method of assessment has two potential advantages. First, it aims to assess teachers in a way that is consistent with their normal activities by having the teachers mark student work. Second, it is likely to be received more positively by teachers than if they were asked to sit the same test as their students. All questions on the teacher test were based on common items from the primary curricula of each country.

The teacher content knowledge score measures the number of correct answers out of the total number of questions in an assessment (literacy for language and numeracy for mathematics teachers). The literacy content knowledge score measures the percentage of questions a teacher answered correctly out of 20 questions, and the mathematics content knowledge score measures the percentage of correct answers out of 23 questions. Table 6 presents a summary of the content knowledge assessment results for teachers who were both assessed (teacher content knowledge assessment) and observed (classroom observation). Table A5 in Annex A shows the same information by education systems.

Table 6: Teacher Content Knowledge

Characteristic	Average	Min	Max
Number of teachers (content)	164.9	58 (Pakistan ICT)	246 (Niger)
Content knowledge	53.9	37.9 (Sierra Leone)	69.8 (Pakistan Punjab)
Literacy			
Number of teachers (literacy)	93.1	28 (Pakistan ICT)	147 (Niger)
Literacy content knowledge	53.6	39.6 (Sierra Leone)	68.9 (Peru)
Mathematics			
Number of teachers (mathematics)	72	31 (Pakistan ICT)	133 (Ethiopia)
Mathematics content knowledge	54.2	31.7 (Pakistan ICT)	84.5 (Jordan)

Sample is restricted to teachers selected for the classroom observation module

On average, teachers performed similarly on both subjects, although some systems did show significant differences between the two: Jordan (with a difference of 31 percentage points), Ethiopia (21 percentage points), Rwanda (20 percentage points), and Islamabad Capital Territory (18 percentage points). Across education systems, teachers answered 54% of the questions correctly, with a low of 38% in Sierra Leone and a high of 70% in Pakistan Punjab.

Learning through Fostering Student Engagement with Learning Remains an Untapped Strategy

Engaging children is an important element of teaching and learning. This has led to increasing interest in ways of supporting children in becoming engaged learners through playful and child-centered practices. Learning through play can support the acquisition of both content (e.g., mathematics) and learning-to-learn skills (e.g., executive function). To assess the extent to which these types of teacher-student interactions take place inside the classroom, we leverage the recently developed PLAY tool.

The PLAY toolkit measures how well adults support students' engagement in learning. The toolkit includes 4 main constructs: support for exploration, support for agency, support for personal and social connection, and support for emotional climate. Each construct includes 5 to 7 practices, for a total of 25 practices. Each practice is scored as "Not observed (0)," "Low quality (1)," and "High quality (2)." We implemented the PLAY instrument in three countries: Ethiopia, Sierra Leone, and Peru. Through the GEPD classroom observation module, we collected classroom observation videos from these 3 countries. These videos were coded using the Teach instrument. We use the PLAY instrument to code the same videos.

Table 7 shows the average score (on a scale of 0-2) for each item in the PLAY instrument, as well as the average score for each construct. Construct scores are the mean of every item within that construct. Overall, support for emotional climate—with items like the use of positive discipline and rewards to support a positive emotional climate, inclusion of children with diverse backgrounds and learning needs, as well as the use of a mode of instruction that is explicitly upbeat—scores relatively higher across systems. The relatively higher performance on these aspects aligns well with the data coming from Teach, where aspects related to creating a supportive learning environment appeared to be more widespread than the rest. Teaching practices associated with support for exploration seem to also be present in many classrooms, particularly in Sierra Leone and Peru. On the other hand, supporting the agency of the students is less common across the 3 countries, signaling how teacher-centric methods continue to be prevalent in low- and middle-income countries, as also illustrated with Teach data.

Table 7: The Use of Teaching Practices to Support Student Engagement in Peru, Sierra Leone, and Ethiopia

	All	Peru	Sierra Leone	Ethiopia
Number of observations	725	183	256	286
PLAY score	.37	.39	.49	.22
	(.02)	(.03)	(.04)	(.01)
	[.25]	[.18]	[.3]	[.16]
Support for Exploration	.47	.48	.67	.26
	(.03)	(.06)	(.06)	(.03)
	[.4]	[.36]	[.44]	[.29]
Teacher connects concepts in the lesson to students' backgrounds, or life outside the classroom	.43	.81	.4	.09
	(.06)	(.15)	(.08)	(.02)
	[.71]	[.84]	[.66]	[.32]
Teacher gives student(s) a chance to try or explore something first before being shown how	.43	.05	.8	.46
	(.05)	(.02)	(.12)	(.06)
	[.69]	[.28]	[.82]	[.62]
Teacher promotes students thinking by themselves	.46	.26	.84	.28

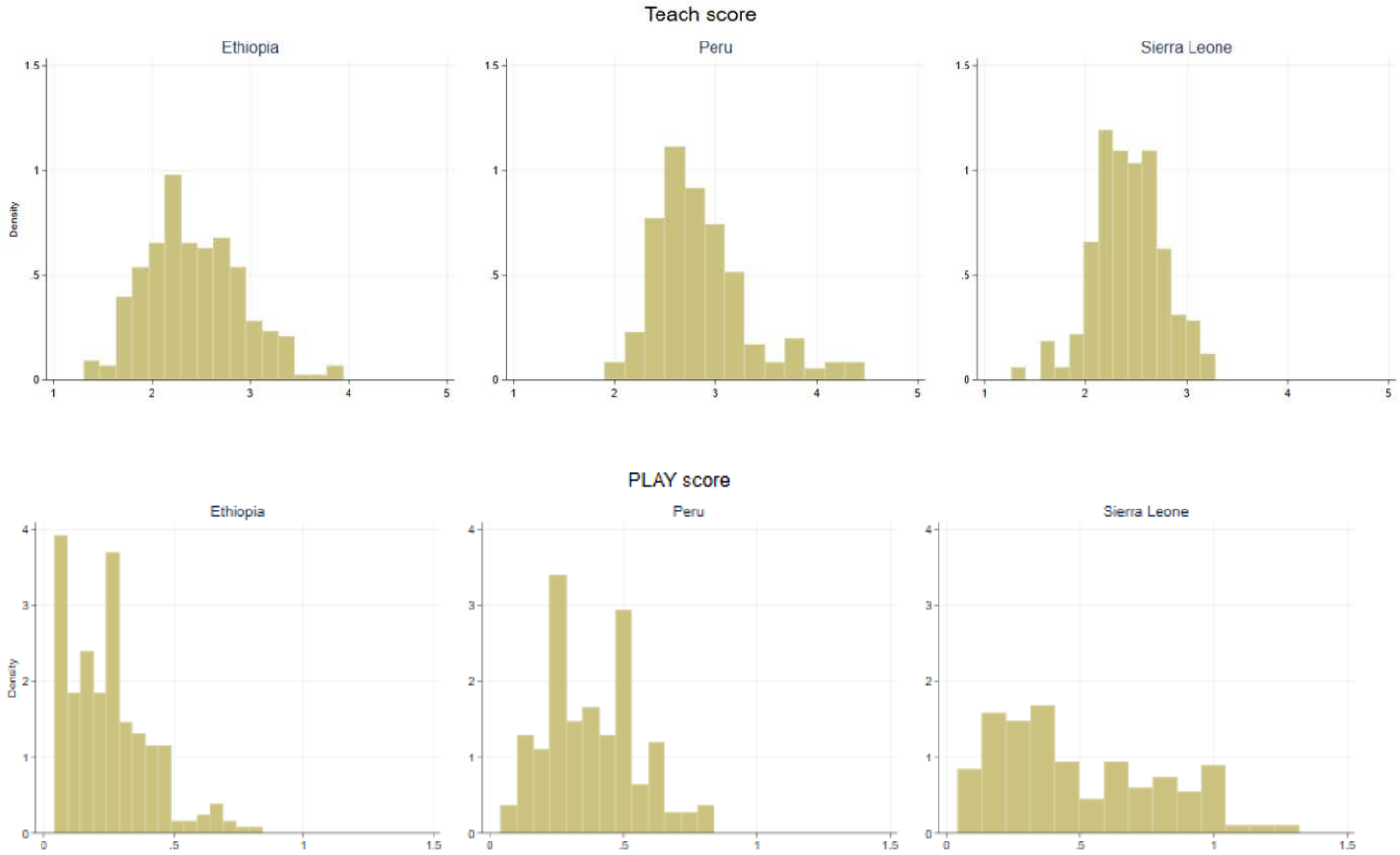
	(.06)	(.07)	(.15)	(.05)
	[.73]	[.5]	[.94]	[.51]
Teacher gives hint or suggestion to encourage students to continue to get to the answer/ explore	.41	.63	.23	.38
	(.05)	(.12)	(.06)	(.05)
	[.66]	[.76]	[.5]	[.64]
Teacher asks questions to generate explanations/reasons	.53	.73	.71	.17
	(.07)	(.16)	(.13)	(.05)
	[.75]	[.82]	[.8]	[.43]
Teacher uses multiple methods to help students learn about a concept	.53	.4	1.04	.16
	(.05)	(.08)	(.1)	(.05)
	[.7]	[.54]	[.76]	[.44]
Support for Agency	.21	.24	.29	.11
	(.02)	(.04)	(.04)	(.01)
	[.25]	[.23]	[.33]	[.13]
Students try different solutions (iteration)	.11	0	.3	.03
	(.04)	(0)	(.1)	(.01)
	[.41]	[.03]	[.65]	[.18]
Students practice a skill introduced by the teacher	.62	.71	.5	.65
	(.07)	(.14)	(.1)	(.07)
	[.74]	[.67]	[.78]	[.76]
Students create something connected to the current lesson	.13	.22	.18	0
	(.04)	(.08)	(.06)	(0)
	[.38]	[.44]	[.47]	[.01]
Students choose WHO plays each role in an activity	.06	.08	.07	.04
	(.02)	(.04)	(.03)	(.03)
	[.28]	[.37]	[.27]	[.19]
Students decide the WHAT or the HOW to do an academic task	.09	.21	.04	.01
	(.03)	(.09)	(.03)	(.01)
	[.33]	[.47]	[.28]	[.08]
Teacher uses student ideas or examples in instruction	.13	.06	.33	.01
	(.03)	(.04)	(.08)	(.01)
	[.43]	[.24]	[.65]	[.1]
Teacher asks open-ended questions or prompts for students to share opinions or preferences	.36	.42	.62	.04
	(.06)	(.13)	(.09)	(.01)
	[.67]	[.72]	[.79]	[.2]
Support for Personal and Social Connection	.23	.37	.28	.06
	(.03)	(.05)	(.04)	(.01)
	[.29]	[.28]	[.34]	[.12]
Teacher has students work together on instructional activity towards a common instructional goal	.48	.93	.44	.07
	(.08)	(.17)	(.09)	(.03)
	[.78]	[.95]	[.67]	[.33]
Teacher discusses or otherwise creates a sense of student/class togetherness	.33	.39	.47	.13
	(.06)	(.12)	(.1)	(.03)
	[.61]	[.65]	[.73]	[.33]
Teacher invites students to share thoughts or personal experience about themselves	.14	.33	.03	.07
	(.05)	(.12)	(.02)	(.04)
	[.43]	[.61]	[.22]	[.28]
Teacher connects student personal interests to learning activities to help them engage	.1	.1	.19	.01
	(.03)	(.08)	(.06)	(.01)
	[.38]	[.3]	[.56]	[.17]
Teacher sets up a task that involves lively social interaction between students	.12	.1	.26	0
	(.03)	(.06)	(.07)	(.)
	[.4]	[.41]	[.54]	[0]
Support for Emotional Climate	.53	.48	.69	.42
	(.02)	(.04)	(.04)	(.02)
	[.29]	[.22]	[.32]	[.26]
Teacher uses a mode of instruction that is explicitly upbeat	.85	.84	1.09	.62
	(.08)	(.17)	(.14)	(.1)
	[.91]	[.89]	[.93]	[.84]
Teacher responds to students' emotional needs	.09	.02	.08	.18
	(.02)	(.01)	(.03)	(.05)
	[.33]	[.17]	[.27]	[.47]
Teacher includes students who did not volunteer to answer	.52	.11	1.12	.33
	(.06)	(.06)	(.11)	(.05)
	[.77]	[.36]	[.84]	[.61]
Teacher makes students know that it is okay to participate even if their answer is wrong	.35	.22	.68	.16
	(.05)	(.07)	(.11)	(.04)
	[.63]	[.48]	[.79]	[.42]
Teacher uses positive discipline and rewards to support a positive emotional climate	.92	1.18	.92	.67
	(.08)	(.15)	(.14)	(.08)
	[.86]	[.77]	[.94]	[.8]

Teacher is inclusive of children with diverse backgrounds and learning needs	.97	1	.92	.98
	(.03)	(.)	(.04)	(.06)
	[.29]	[0]	[.27]	[.41]
The teacher explains a student's actions, intentions, and/or feelings to other students	.02	0	.04	.01
	(.01)	(0)	(.02)	(.01)
	[.13]	[.03]	[.2]	[.09]

Table shows average with standard error in parentheses and standard deviations in brackets. Values in the overall column are the unweighted average across countries. Values in country columns are weighted.

The comparison of Teach and PLAY scores for these 3 countries—Ethiopia, Peru, and Sierra Leone—offers insights into the prevalence of practices to foster student engagement with learning in classrooms. Figure 1 shows the distribution of Teach and PLAY scores in the 3 countries. PLAY scores tend to skew lower relative to its scale than Teach scores do in Ethiopia and Peru, signaling that practices measured through Teach may be more commonly observed in classrooms than those captured in the PLAY rubric. Furthermore, PLAY scores appear more variable than Teach scores in one of the 3 countries (Sierra Leone). This could be because pedagogies to foster student engagement, including through playful and child-centered activities, are not prescribed by the system, resulting in more idiosyncratic variation based on teachers' interests and abilities.

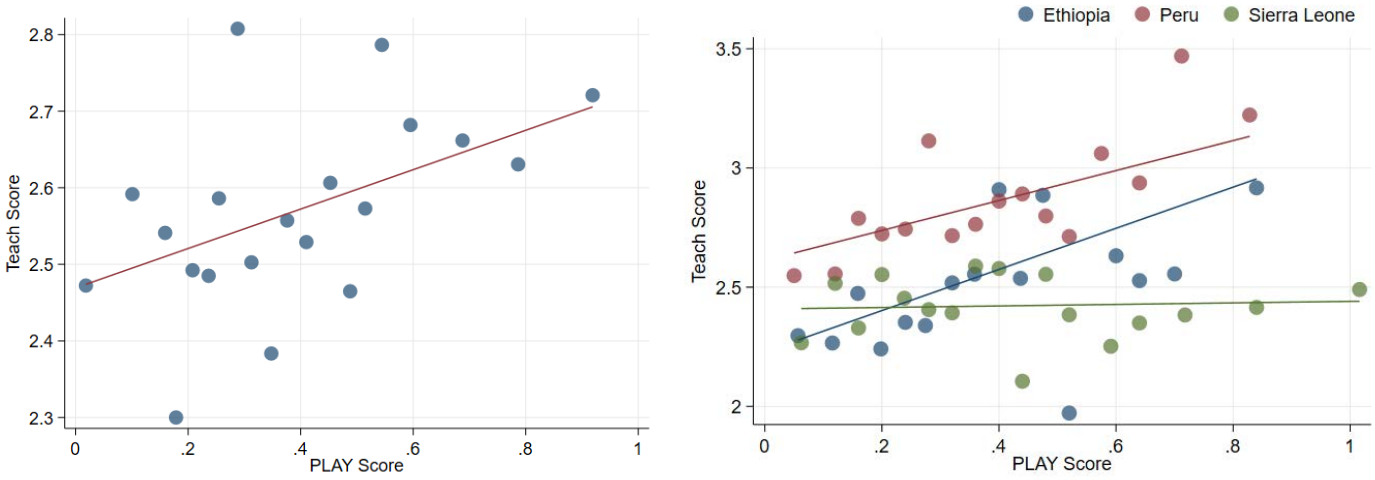
Figure 1: Histogram of Teach and PLAY Scores by Country



As previously mentioned, there appears to be a correlation between Teach and PLAY scores, signaling a correspondence between the 2 tools (as in, they may capture similar skills) as well as a correlation between the

skills they capture (as in, even though they capture some different skills, these skills tend to correlate with each other). Figure 2 more explicitly presents the relationship between the 2 tools by showing a binned scatterplot of PLAY and Teach scores. In both Ethiopia and Peru, the Teach and PLAY scores are positively correlated, while there is no significant relationship in Sierra Leone. A detailed comparison of the tools, making use of the Delphi method, shows that 9 (or 36%) of the PLAY items fully match Teach Primary items (Bazaldua, Carter & Gregory, forthcoming). The rest of the items, on both tools, do not have a corresponding match, which highlights the PLAY tool's ability to capture other aspects of teaching conducive to fostering student engagement with instruction and learning.

Figure 2: Relationship between Teach and PLAY



Teachers Lack Adequate Support Inside and Outside the Classroom

The issues related to teaching quality extend well beyond schools. Challenges with the quality of teaching often indicate more profound issues within the system. These are caused by policies that are either poorly designed or simply not implemented effectively, failing to foster learning for every child.

Across countries, the policies that consistently require improvement relate to the support systems available to teachers. While policy frameworks tend to be comprehensive, they are often not implemented effectively or fail to reach every teacher. As demonstrated in Table A6 of Appendix A, teaching practicums during pre-service remain rare, with most education systems having less than 50% of their teachers having benefited from one (except for Gabon, Niger, Peru, and Sierra Leone). In-service training often does not reach every teacher, with only 3 education systems (ICT, Peru, and Rwanda) providing in-service training to at least 50% of teachers within a 12-month period. Even when in-service training does take place, it is often not practical. Across all education systems, 85% of teachers report their last in-service training had no in-classroom component.

Extensive evidence has also highlighted the importance of supporting teachers through effective instructional leadership. A key component of effective instructional leadership is the ability to detect and assist teachers and students who are facing challenges. Utilizing classroom observations is a widely recognized and beneficial approach for assessing teachers' capabilities. This method allows leaders to pinpoint specific areas of strength and areas needing development, thereby offering precise guidance for enhancing teaching methods (Beisiegel, Mitchell, & Hill 2018). To an extent, we do see practices of instructional leadership being mainstreamed through

education systems. Across all the education systems, the average share of teachers reporting that their classroom was observed over the past 12 months was 75%. In most cases, the classroom observation was recent, with the average being 2 months before the interview. There are, however, areas that signal the need to continue to improve instructional leadership practices. The lack of follow-up after the classroom observation is one. Across systems, the average shares of teachers reporting having an opportunity to discuss the results of the observation and receiving feedback after the observation were only 58% and 54%, respectively.

Other aspects such as the extent to which teachers are evaluated, the incentives for good performance, the intrinsic motivation of teachers, or even the standards considered when it comes to recruiting new teachers remain challenges in some education systems. Please refer to Table 8 for a summary across all systems.

Table 8: Policy Levers for Teaching Quality

Characteristic	Average	Min	Max
Factors considered during the recruitment:			
Completed required coursework	25.9	8.3 (Gabon)	52.6 (Ethiopia)
Achieved special educational qualification	59.2	33 (Pakistan ICT)	85.5 (Jordan)
Graduated from tertiary education degree program	25.3	11.2 (Ethiopia)	85.8 (Rwanda)
Graduated from tertiary program designed to prepare teachers	27.9	12.5 (Pakistan Sindh)	68.3 (Pakistan ICT)
Passed a subject content knowledge written test	42.4	14.6 (Sierra Leone)	78.1 (Rwanda)
Passed an interview stage assessment	41.8	9.8 (Niger)	72 (Pakistan Punjab)
Had a minimum amount of practical professional experience	17.5	4.6 (Pakistan ICT)	33.3 (Gabon)
Passed a practical assessment conducted by a supervisor	14.4	1.7 (Pakistan Punjab)	40.3 (Peru)
Conduct during mockup class	6.6	0 (Pakistan Balochistan, Pakistan Sindh)	27.6 (Rwanda)
Teacher training:			
Probationary period for new teachers	46.4	12 (Pakistan Sindh)	81.1 (Pakistan ICT)
Participated in induction/mentorship program	44.7	9.6 (Pakistan Balochistan)	64.9 (Gabon)
Teaching practicum under pre-service training	33.4	6.6 (Pakistan KP)	63.4 (Peru)
Hours per day spent teaching during teaching practicum	1.3	.1 (Pakistan ICT)	3.1 (Gabon)
Participated in in-service training	39.2	5.6 (Pakistan Balochistan)	92.7 (Peru)
Days in-service training lasted	3.3	.9 (Gabon)	10.4 (Peru)
No in service training done in classroom	84.8	40.1 (Peru)	100 (Pakistan Balochistan)
Less than 50% in service training done in classroom	9.2	0 (Pakistan Balochistan)	28.7 (Chad)
Over 50% in service training done in classroom	6.1	0 (Pakistan Balochistan, Pakistan ICT, Pakistan Punjab)	37.6 (Peru)
Instructional leadership:			
Classroom has been observed	74.9	36 (Pakistan Sindh)	96.2 (Pakistan ICT)
Months since last classroom observation	1.9	.7 (Pakistan KP)	3.4 (Ethiopia)
Discussed results of observation	58.4	17.8 (Pakistan Balochistan)	83.2 (Gabon)
Observer provided feedback	53.8	12.4 (Pakistan Balochistan)	79.7 (Gabon)
Had lesson plans for last week	73.7	38.3 (Pakistan Balochistan)	99.3 (Peru)

Discussed lesson plans	44.5	16.5 (Pakistan Balochistan)	86.8 (Peru)
Evaluation:			
Evaluated on pedagogical skill and content knowledge	53.8	20.1 (Pakistan Balochistan)	89.6 (Rwanda)
Financial consequences to negative evaluation	81.6	63.2 (Jordan)	99.2 (Peru)
Financial consequences to positive evaluation	70.2	30.2 (Rwanda)	94.3 (Pakistan Punjab)
Incentives:			
Meritocracy for advancement	50.4	17.7 (Pakistan Sindh)	94.2 (Sierra Leone)
Received bonus during last academic year	9.9	0 (Pakistan Punjab, Pakistan Sindh)	42.1 (Rwanda)
Salary was delayed during last academic year	16.2	.9 (Pakistan Punjab)	37 (Pakistan ICT)
Intrinsic motivation:			
Motivation - always wanted to be a teacher and likes teaching	55.7	29.1 (Pakistan ICT)	82.9 (Pakistan Sindh)
Motivation - teaching offers steady career path	7.4	.6 (Peru)	16.8 (Sierra Leone)
Motivation - benefit the socially disadvantaged	17.8	2.2 (Jordan)	49.9 (Peru)

What Do We Know About the Correlation Between Teaching and Learning?

To explore the relationship between teaching and student learning, we estimate how well the test scores of Grade 4 students are predicted by their teachers' pedagogical skill (*Teach* score) and content knowledge. Our sample for these estimations is restricted to Grade 4 students who are taught by a teacher selected for the classroom observation exercise. As noted above, one teacher in each school was selected for this Teach classroom observation. Our sample therefore includes 2,686 teachers selected to complete the classroom observation module and 50,392 students across the 13 systems. When we include teacher content knowledge, our sample consists of 2,144 teachers who completed both the classroom observation module and the content knowledge assessment.

We use 2 different approaches to model selection. First, we estimate a core model with key variables based on mechanisms identified in existing literature (also our preferred specification). Second, because the GEPD generates so many indicators across a wide range of domains (service delivery, policy, and political/bureaucratic), we use three machine learning models for variable selection. This approach serves as a robustness check on the results from the core model, as in Filmer, Nahata, and Sabarwal (2021). Below, we describe each of these model specifications in turn. See Figure 3 for further details.

Core model

Our core model specification is as follows:

$$Y_{ijks} = \beta_1(I\ V_{jkt}) + \beta_2(W_{ijk}) + \beta_3(X_{jk}) + \beta_4(Z_k) + \varepsilon \qquad (1)$$

where Y_{ijks} is the test score of Grade 4 student i in school j in country k in subject s, where s is the *student knowledge score, literacy knowledge score,* or *mathematics knowledge score*. Student knowledge score is the average of a student's mathematics and literacy scores; literacy knowledge score is the percentage of questions a student got correct out of 6 questions; and mathematics knowledge score is the percentage of correct answers a student got out of 15 questions. V_{jkt} is a measure of teaching quality in school j in country k. We consider 3 variations for teaching quality, denoted by t = *{teach score, teacher content knowledge,* or *teach score* and *teacher*

content knowledge together*}*. Teach score (1-5) is the average score across all elements in the Teach instrument. Teacher content knowledge (0-100) measures the percentage of questions answered correctly, out of 20 questions for literacy assessments and 23 for numeracy assessments.

W_{ijk} is a vector of student characteristics and includes student's age and gender. X_{jk} is a vector of teacher characteristics including gender, teaching experience in years, and the teacher's highest level of education. Finally, Z_k is a vector of school controls, including rurality (dummy taking the value of 1 if school is in a rural area), number of students enrolled, average teacher absence rate, infrastructure index, and a school-level wealth index. The infrastructure index (0 to 5) measures whether a school has access to drinking water, functioning toilet, functioning internet, and functioning electricity in the classroom and whether the school is accessible to students with disabilities, with equal weight given to each type of access.

Extended model

While the core model includes variables highlighted by the literature as correlates of student learning, the extended model draws on the full set of GEPD variables. To reduce the dimensionality of this very large set of variables, we use 3 different machine learning approaches to identify the variables most correlated with the Grade 4 test scores. The 3 algorithms we used are the Conditional Inference Forest (CIF), Random Forest (RF), and LASSO algorithms. This approach is similar to Filmer, Nahata, and Sabarwal (2021), who apply the CIF and LASSO algorithms to identify teacher covariates that are predictive of student learning gains in Tanzania. The authors find that the machine learning algorithms outperform standard OLS regressions by 14%-24% on out-of-sample prediction. Additionally, they use the results from the CIF and LASSO to choose regressors in an OLS regression of student knowledge scores on several covariates. The authors note that the CIF algorithm has better performance identifying discrete indicators as important covariates than the RF algorithm, and so more weight was placed on the results of the CIF algorithm in choosing variables for the extended model.

When implementing the algorithms, the sample is randomly split into a training sample composed of 70% of the observations and a testing sample composed of the remaining 30% of the sample, which is used to validate the performance of the algorithm on data not used for the training to avoid problems of over-fitting the data.

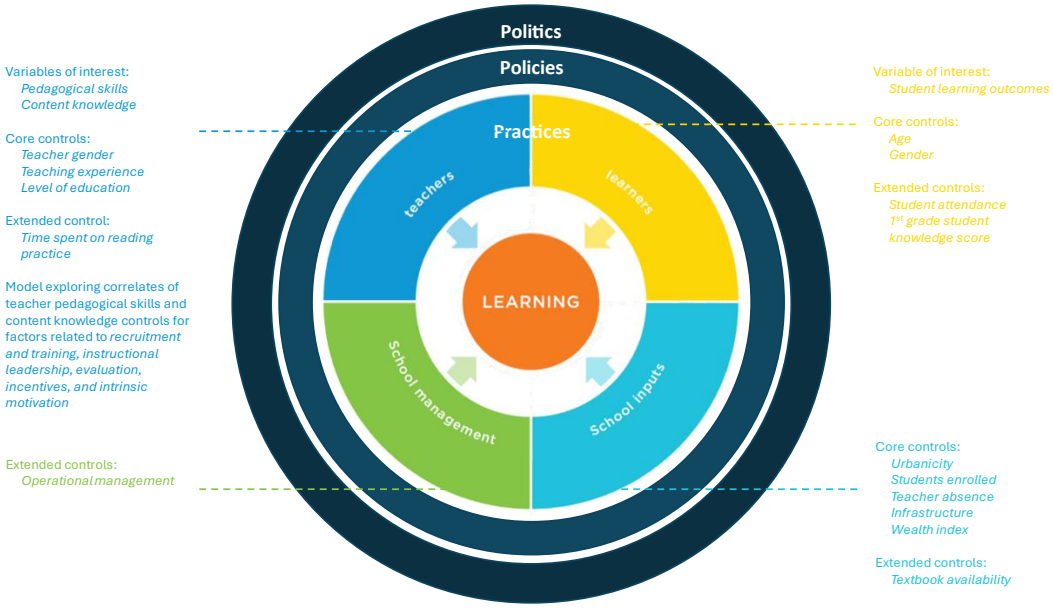
Similar to Filmer, Nahata, and Sabarwal (2021), we find that the machine learning models have significantly better predictive power for student learning compared to OLS. The OLS estimator in a regression of Grade 4 student knowledge on the full set of GEPD variables has an R^2 value of 0.349, while the RF algorithm has an R^2 value of 0.589 and the CIF algorithm has a value of 0.587, nearly two-thirds higher than the OLS estimator. For the regression of the Teach scores on the GEPD variables, OLS has an R^2 value of 0.399, while the RF algorithm has an R^2 value of 0.99 and the CIF has an R^2 value of 0.797.

Based on the results of these machine learning models, we selected additional variables as controls and included them in our extended model. The extended model uses the same specification as in equation (1), but with the additional variables identified as important by the machine learning models. Specifically, we add student attendance in our vector W_{ijk}; time (in minutes) spent on reading practice in vector X_{jk}; and operational management, availability of textbooks, and the average Grade 1 student knowledge score in the vector Z_k.[8] Operational management is the principal's or head teacher's score on responses to two vignettes on school management. The variable importance rankings for these indicators from the RF and CIF machine learning models are available in Table A7 of Appendix A.

[8] For the models predicting student learning, student attendance was the 11th most important in the RF and 28th most important in the CIF model out of 138 variables considered. Time spent on reading practices was the 13th most important in the RF and 71st in the CIF.

Figure 3 presents the conceptual framework underlying the estimation models. It maps the key variables used in our models onto the GEPD framework on the drivers of foundational learning outcomes.

Figure 3: Conceptual Framework



Better Teacher Pedagogical Practices Are Predictive of Higher Student Learning Outcomes

This section explores how well teacher pedagogical skill and knowledge predict Grade 4 student test scores, focusing on the coefficients on the teacher skills variables. Table 9 summarizes the pooled results across all education systems. It gives the coefficient on each teacher skills variable, with the different columns representing different dependent variables and specifications. The dependent variables for the different regressions are overall student test scores (in Columns 1, 2, and 3), student literacy scores (Columns 4, 5, and 6), and student mathematics scores (Columns 7, 8, and 9). In each pair of columns, the first column shows the coefficient with no controls other than country fixed effects, the second includes the full set of controls in the core model, while the third column includes the full set of controls in the extended model. Finally, the different panels use different teacher skill measures as independent variables: Panel A uses teacher pedagogical skill as measured by the Teach score (1-5), Panel B uses teacher content knowledge, Panel C includes both the Teach score and content knowledge, Panel D includes pedagogical skill sub-scores. Standard errors are clustered at the school level.

The results of the core model consistently indicate that higher teacher skills and knowledge are associated with higher test scores for their students:

- *Pedagogical skill:* In all specifications in Panel A, the teacher's Teach score is positively correlated with student scores. Without controls, a 1-point increase in Teach score predicts an increase in overall student test scores by 3.31 points, with slightly larger effects for literacy. The result is significant at the 1% level in each case. Adding the full range of controls—student, teacher, and school characteristics—attenuates the coefficient, as we would expect. Nevertheless, a 1-point increase in Teach still predicts an increase of 1.91 points in overall test scores, 2.09 points in literacy, and 1.72 points in mathematics. Expressed another way, a 1-SD increase in Teach predicts an increase of 1.34 points in overall test scores, 1.46 points in

literacy, and 1.20 points in math. Adding the full range of controls eliminates the significance of the results.

- *Teacher content knowledge:* Teacher content knowledge (Panel B) is also robustly positively correlated with student knowledge (overall, literacy, and mathematics). Without controls, each 1-point increase in content knowledge is associated with an increase of around 0.10 points in test scores. With full controls added, the coefficient falls to about 0.08. Since content knowledge is scored 0-100, a more relevant way to express the relationship is to say that a 10-point increase in teachers' content knowledge predicts a nearly 1-point increase in student test scores (without controls) or a 0.8-point increase (with full controls). The corresponding increase for literacy and math scores is 0.68 points and 0.95 points respectively. In other words, a 1-SD increase in content knowledge predicts an increase of 1.83 points in overall test scores, 1.52 points in literacy, and 2.12 points in math, respectively. All coefficients are highly significant.[9]
- *Pedagogical skill and content knowledge together:* When the two independent variables are both included in the regressions, the magnitude and significance of the Teach coefficients are slightly reduced. Nevertheless, even with the full set of controls, a 1-point increase in Teach score still predicts an increase of 1.58 points in overall student scores, 1.71 points on the literacy test, and 1.45 points on the math test. The coefficients on content knowledge barely change in magnitude with the inclusion of the Teach variable, and they remain significant at the 1% level, even with full controls.

The relationship between Teach scores and student learning is even stronger when we focus on teachers' skills in the subject on which the student is tested. Table 9(b) shows the effects of this restriction. For example, only some teachers teach language only, not math; these teachers are included only in the regressions presented in Columns 4, 5, and 6 of Table 9(b), which predict literacy scores. Similarly, some other teachers teach only math, and so they are included only in Columns 7, 8, and 9. Teachers who teach both subjects are included throughout. Results from restricting the sample to teachers who teach language (either only language or both subjects) show that a 1-point increase in Teach predicts an increase of 3.13 points in literacy. Similarly, results from restricting the sample to teachers who teach either only math or both subjects show that a 1-point increase in Teach predicts an increase of 1.82 points in math. Expressed another way, a 1-SD increase in Teach predicts an increase of around 2.2 points in literacy and 1.3 points in math. Looking at teacher content knowledge, a 1-point increase in content knowledge for teachers who teach language is associated with an increase of around 0.087 points in student literacy scores. Similarly, a 1-point increase in content knowledge for teachers who teach math is associated with an increase of around 0.099 points in student math scores. Expressed another way, a 1-SD increase in teacher content knowledge predicts an increase of around 1.94 points in literacy and 2.21 points in math. With both Teach scores and content knowledge are added to the restricted model, the magnitude and significance of Teach is lower but still meaningful and significant in our preferred specification.

Table C2 in Appendix C presents the country-by-country results. Note that the relationship between pedagogical skills, content knowledge, and student learning outcomes varies by country, which may be driven by smaller sample sizes or contextual factors at the country level. This underscores the importance of further research on the role of pedagogical skills and content knowledge at the country level.

[9] As noted above, these regressions include only the content knowledge of the one teacher in each school whose teaching was recorded for the classroom observation (typically the teacher of the Grade 4 students who were tested). When we also include the average content knowledge score of the other teachers who were tested in each school (see Appendix C, Table C3), the coefficient on content knowledge declines in magnitude and becomes less significant for literacy. This is what would be expected, if the single teacher's score had partly been proxying for the knowledge of all teachers in the school (including others that the students have studied with in the past). The coefficients on Teach remain significant predictors of both literacy and math scores. For Panel C perhaps the most relevant, because Teach and content knowledge enter together—the coefficient on Teach for the core model persists across all outcomes.

Table 9: Regression Results for Core and Extended Models

(a) Full Sample

	(1) Student knowledge	(2) Student knowledge	(3) Student knowledge	(4) Literacy student knowledge	(5) Literacy student knowledge	(6) Literacy student knowledge	(7) Math student knowledge	(8) Math student knowledge	(9) Math student knowledge
Panel A: Teacher pedagogical skill									
Teach score	3.31***	1.91***	1.06	3.57***	2.09***	1.11	3.05***	1.72**	1.00
	(0.53)	(0.66)	(0.65)	(0.58)	(0.73)	(0.73)	(0.59)	(0.74)	(0.72)
Observations	50104	38912	34630	50108	38912	34630	50106	38912	34630
R^2	0.220	0.273	0.316	0.289	0.332	0.358	0.098	0.135	0.174
Panel B: Teacher content knowledge									
Content knowledge	0.095***	0.082***	0.067***	0.087***	0.068***	0.053***	0.10***	0.095***	0.082***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	39785	36198	32234	39785	36198	32234	39785	36198	32234
R^2	0.219	0.270	0.311	0.294	0.340	0.364	0.094	0.126	0.163
Panel C: Pedagogical skill and content knowledge									
Teach score	2.63***	1.58**	0.77	3.06***	1.71**	0.82	2.20***	1.45*	0.72
	(0.62)	(0.68)	(0.66)	(0.67)	(0.73)	(0.73)	(0.69)	(0.76)	(0.74)
Content knowledge	0.090***	0.080***	0.066***	0.082***	0.067***	0.051***	0.099***	0.093***	0.080***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	39665	36079	32161	39665	36079	32161	39665	36079	32161
R^2	0.224	0.272	0.312	0.297	0.341	0.364	0.098	0.130	0.166
Panel D: Pedagogical skill sub-scores									
Classroom culture	0.86	0.79	0.0044	0.82	0.79	-0.12	0.90	0.78	0.13
	(0.60)	(0.68)	(0.68)	(0.69)	(0.77)	(0.78)	(0.66)	(0.76)	(0.76)
Instruction	0.71	0.17	-0.028	1.00	0.27	-0.085	0.42	0.068	0.029
	(0.63)	(0.68)	(0.68)	(0.71)	(0.76)	(0.78)	(0.71)	(0.81)	(0.81)
Socio emotional skills	1.75***	0.99	1.10	1.74**	1.06	1.34	1.76**	0.92	0.86
	(0.63)	(0.69)	(0.69)	(0.74)	(0.80)	(0.82)	(0.69)	(0.77)	(0.77)
Observations	50104	38912	34630	50108	38912	34630	50106	38912	34630
R^2	0.221	0.273	0.316	0.289	0.332	0.358	0.098	0.136	0.174
School characteristics	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Teacher characteristics	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Student characteristics	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Extended controls	No	No	Yes	No	No	Yes	No	No	Yes
Survey time	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table presents point estimates from regressing student learning (student knowledge score) on dimensions of teaching quality. In the extended model (Column 3, 6, and 9), we add variables as controls based on the outcome of the CIF, random forest and lasso models. Columns 1, 2, and 3 refer to overall student knowledge, columns 4, 5 and 6 refer to student knowledge in literacy, and columns 7, 8 and 9 refer to student knowledge in mathematics. In panel A, teacher pedagogical skill is measured by the TEACH score (1-5), an average of all items in the TEACH instrument. In panel B, teacher content knowledge measures the percentage of correct questions on the teacher content knowledge test (language for language teachers and mathematics for mathematics teachers), ranging from 0 to 100. Panel C includes both pedagogical skill and content knowledge, as described

earlier, while panel D includes pedagogical skills sub-scores. In columns 2,4 and 6, the model controls for a vector of student, teacher, and school characteristics. Student controls include age and gender. Teacher controls include gender, teaching experience in years, and teacher's level of education. School controls include rurality, number of students enrolled, average teacher absence rate, infrastructure index, and a school-level wealth index. The infrastructure index measures whether a school has access to drinking water, functioning toilet, functioning internet, and functioning electricity in the classroom, and whether the school is accessible to students with disabilities. These models also control for survey time, the number of months from start of academic year to implementation of survey instrument. All models include country fixed effects. Standard errors clustered by school in parentheses. * p< 0.1, ** p< 0.05, *** p< 0.01

(b) Restricted Sample Based on Teacher Subject(s) Taught

	(1) Student knowledge	(2) Student knowledge	(3) Student knowledge	(4) Literacy student knowledge	(5) Literacy student knowledge	(6) Literacy student knowledge	(7) Math student knowledge	(8) Math student knowledge	(9) Math student knowledge
Panel A: Teacher pedagogical skill									
Teach score	4.41***	2.86***	1.74**	4.74***	3.13***	2.08**	3.13***	1.82**	1.00
	(0.78)	(0.90)	(0.85)	(0.69)	(0.82)	(0.81)	(0.75)	(0.85)	(0.83)
Observations	29529	25119	22914	40488	33410	31041	36686	30543	26475
R^2	0.256	0.321	0.366	0.298	0.346	0.366	0.102	0.147	0.191
Panel B: Teacher content knowledge									
Content knowledge	0.13***	0.11***	0.092***	0.11***	0.087***	0.065***	0.12***	0.099***	0.088***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	26196	24328	22175	33703	31026	28769	32091	29422	25612
R^2	0.262	0.323	0.367	0.313	0.357	0.375	0.100	0.141	0.186
Panel C: Pedagogical skill and content knowledge									
Teach score	3.56***	2.44***	1.37	3.97***	2.65***	1.71**	2.38***	1.67*	0.79
	(0.82)	(0.90)	(0.85)	(0.74)	(0.81)	(0.80)	(0.79)	(0.86)	(0.84)
Content knowledge	0.12***	0.11***	0.090***	0.099***	0.084***	0.062***	0.11***	0.098***	0.087***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	26101	24234	22102	33608	30932	28696	31971	29303	25539
R^2	0.270	0.327	0.369	0.318	0.359	0.375	0.105	0.146	0.190
Panel D: Pedagogical skill sub-scores									
Classroom culture	1.67**	1.46*	0.16	1.44*	1.18	0.10	1.06	1.12	0.25
	(0.83)	(0.84)	(0.82)	(0.79)	(0.83)	(0.83)	(0.82)	(0.88)	(0.89)
Instruction	0.67	-0.11	0.10	1.39*	0.76	0.59	-0.18	-0.76	-0.67
	(0.82)	(0.81)	(0.81)	(0.76)	(0.78)	(0.80)	(0.89)	(0.93)	(0.93)
Socio emotional skills	2.19**	1.69*	1.55*	1.93**	1.23	1.39	2.36***	1.60*	1.51*
	(0.86)	(0.88)	(0.86)	(0.81)	(0.86)	(0.87)	(0.87)	(0.89)	(0.88)
Observations	29529	25119	22914	40488	33410	31041	36686	30543	26475
R^2	0.256	0.321	0.367	0.298	0.346	0.366	0.103	0.148	0.192
School characteristics	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Teacher characteristics	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Student characteristics	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Extended controls	No	No	Yes	No	No	Yes	No	No	Yes
Survey time	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table presents point estimates from regressing student learning (student knowledge score) on dimensions of teaching quality. In this table, we restrict teachers depending on the subject taught: columns 1, 2, and 3 include only teachers who teach both language and math, columns 4, 5, and 6 include teachers who teach either both language and math or language only, and columns 7, 8, and 9 include teachers who teach either language and math or math only. In the extended model (Column 3, 6, and 9), we add variables as controls based on the outcome of the CIF, random forest and lasso models. Columns 1, 2, and 3 refer to overall student knowledge, columns 4, 5 and 6 refer to student knowledge in literacy, and columns 7, 8 and 9 refer to student knowledge in mathematics. In panel A, teacher pedagogical skill is measured by the TEACH score (1-5), an average of all items in the TEACH instrument. In panel B, teacher content knowledge measures the percentage of correct questions on the teacher content knowledge test (language for language teachers and mathematics for mathematics teachers), ranging from 0 to 100. Panel C includes both pedagogical skill and content knowledge, as described earlier, while panel D includes pedagogical skills sub-scores. In columns 2,4 and 6, the model controls for a vector of student, teacher, and school characteristics. Student controls include age and gender. Teacher controls include gender, teaching experience in years, and teacher's level of education. School controls include rurality, number of students enrolled, average teacher absence rate, infrastructure index, and a school-level wealth index. The infrastructure index measures whether a school has access to drinking water, functioning toilet, functioning internet, and functioning electricity in the classroom, and whether the school is accessible to students with disabilities. These models also control for survey time, the number of months from start of academic year to implementation of survey instrument. All models include country fixed effects. Standard errors clustered by school in parentheses. * p< 0.1, ** p< 0.05, *** p< 0.01

When we compare results across country income levels (see Appendix C, Table C1), we see markedly different results. The Teach scores are much more correlated with student learning in the middle-income countries than in the low-income countries in our sample, especially for literacy:

- For the *low-income countries (LICs)*, which represent around half of the student observations in the sample, the coefficients on the Teach pedagogical score are not significant, and the point estimates for the impact on literacy are lower than for the full group of countries. By contrast, the coefficients on content knowledge are always significant at the 1% level. Their significance persists even when full controls and Teach score are included in the regressions.
- For the *lower-middle-income countries (LMICs),* the coefficients on Teach are larger in magnitude than those for the full group of countries for literacy, and they continue to be significant as in the full group of countries. The coefficients on Teach are especially large for Balochistan (8-point increase in overall scores for a 1-point increase in Teach scores, statistically significant at the 5% level) and KP (6-point increase, statistically significant at the 1% level) provinces of Pakistan (see Appendix C, Table C2).
- Finally, for the *upper-middle-income countries (UMICs)* in the sample, the coefficient on Teach is much larger for literacy. Even with full controls and content knowledge included, a 1-point increase in Teach score predicts a 5-point increase in literacy scores. This is a little less than two times as high as in the LMICs, and much higher than the insignificant coefficient for the LICs. The result remains significant for literacy (0.05 level), despite the much smaller number of observations for this group (only about 10% of the total observations). The coefficient is especially large for Gabon (6-point increase in the overall student test results for a 1-point increase in Teach scores with full controls and content knowledge added in the regression) (see Appendix C, Table C2).

One possible concern with the full-controls specifications of the core model is that by including teacher education as a control, we could be underestimating the effect of teachers' pedagogical skills and content knowledge, because we would expect that more education would lead to higher levels of both. Indeed, results presented below (in Tables 11 and 13) show that teachers' education is positively correlated with both their Teach and content knowledge scores. Our core specification includes teacher education because we recognize that Teach and content knowledge scores will be imperfect measures of what teachers know and can do, and having more education may have improved their teaching and their students' learning in other ways that are hard to measure. As a robustness check, however, we re-run the specifications without the teacher education variables (see Appendix C, Table C4). The change has only minor impacts: without education included, the coefficients on Teach and content knowledge change only slightly.

In the extended-model results, the robustness of the results depends partly on whether we match teachers to student scores based on subject taught. With the full sample (Columns 3, 6, and 9 in Table 9(a)), teachers' content knowledge (Panel B) remains a strongly significant predictor of Grade 4 test scores, in both literacy and mathematics, and the coefficients are slightly lower than those in the core model. The coefficients on pedagogical skill (Panel A), however, diminish in magnitude and lose significance with the addition of the new extended-model variables. When the teachers' content knowledge is added to the regression (Panel C), the coefficients on Teach become insignificant for all outcomes. But in the model linking the Teach scores to student scores based on subject taught—which is the more properly specified model for assessing effects of pedagogical skill—the Teach coefficients remain robustly significant for literacy, even in the extended model results. Moreover, in interpreting these extended-model results, it is important to bear in mind that the additional variables in this specification were identified as correlates by the machine-learning analysis, rather than being derived from a causal conceptual model. Two of the added variables—student attendance and the class time devoted to reading practice—are plausibly endogenous to the dependent variable of learning, potentially biasing the coefficients.

Summary and interpretation

These results suggest that measurable dimensions of teacher knowledge and skills can matter for student learning in low- and middle-income countries. Even with a wide range of controls, the quality of pedagogy (as measured by Teach) is significantly correlated with student learning. In our data, pedagogical skill is more strongly associated with literacy scores than with mathematics scores, and this literacy effect is strongest in the middle-income countries in our sample, especially in the upper-middle-income countries. Teacher content knowledge is also a significant predictor of student learning; the predicted content knowledge impact is at least as strong for mathematics as for literacy, and unlike the pedagogical skill, it is highly significant for the low-income countries.

Some caveats and nuances to these overall results are necessary.

First, the coefficients on the teacher skill variables are modest. Even in the core results (with all controls) in Table 9, a 1-point increase in the Teach pedagogical score (a 1-to-5 scale) is associated with literacy score increases of only 2.09 points out of 100 (and overall score increases of 1.91 points). For content knowledge (a 0-to-100 scale), a 10-point increase predicts a student test score increase of less than 1 point out of 100 (although a highly significant one). Including both teaching practices and content knowledge in the model reduces the magnitude of impact on student scores associated with teaching practices, while those related to content knowledge stay similar.

It is likely, however, that this empirical set-up understates the impact of teachers' skills and knowledge. The regressions correlate test scores of Grade 4 students with the Teach and content knowledge scores of only their current teacher, because we have Teach pedagogical data for only one teacher per school. But Grade 4 students' test scores reflect not only teaching by their current teachers but also teaching in earlier grades. Since we do not have pedagogical data for those earlier grades, the regressions do not capture the overall impacts of pedagogical quality on student learning.[10] Also, even for the current teachers, the Teach instrument and content-knowledge assessment—while better than what has been deployed before across such a range of countries—provide only (somewhat blurry) snapshots of what teachers know and do. Any measurement error will also tend to attenuate the correlations.

Second, while the GEPD sample includes a limited number of countries in each group, the results suggest that patterns in low-income countries may differ from those in middle-income countries. For the low-income countries as a group, Teach scores have a small and statistically insignificant impact on student test scores, whereas teachers' content knowledge does matter. In contrast, in the lower- and upper-middle-income countries in our sample, the coefficients are much larger: a 1-point increase in Teach yields a predicted 2.8- to 5.3-point increase in student literacy scores, correspondingly, with full controls and teacher content knowledge added (see Annex C, Table C1, Panel C).

One way to interpret this apparent stronger Teach effect in middle-income countries is to consider how teacher skills and knowledge interact in promoting student learning. Stated simply, teaching quality is the product of the teacher's knowledge of the material (content knowledge) and her ability to convey it effectively in ways that will help students learn (Teach pedagogical score). When the content knowledge of most teachers is especially low, as in some low-income countries, stronger pedagogical skills may not have much impact. By contrast, once the teacher has at least a basic command of the material—as we would expect in UMICs—pedagogy could become a more important determinant of student learning. These differential LIC-MIC patterns merit more investigation.

[10] As a sensitivity check, we added the school average teacher content knowledge score as a regressor in our regressions (see Table C3, Appendix C), as a proxy for the teacher content knowledge the student experienced in past grades. While this additional control, the results still remain statistically significant.

Teaching Practices to Foster Student Engagement with Learning Are Predictive of Better Student Learning Outcomes

Learning through play can be a powerful tool to help students engage better with the content and thereby improve their numeracy and literacy skills, among others. Using the scores from the PLAY (Playful Learning Across the Years) classroom observation tool, we estimate to what extent teaching practices that focus on supporting student engagement with learning matter for student learning.

Modeling approach

We link the PLAY data with the GEPD teacher level data and estimate the correlation between practices that support student engagement with the instructional content and learning outcomes.

Our model specification is as follows:

$$Y_{ijks} = \beta_1(P_{jkt}) + \beta_2(W_{ijk}) + \beta_3(X_{jk}) + \beta_4(Z_k) + \varepsilon \qquad (2)$$

where Y_{ijks} is the student knowledge score of student *i* in school *j* in country *k* in subject *s = {student knowledge score, literacy knowledge score, mathematics knowledge score}*. P_{jkt} is the PLAY score of a teacher in school *j* in country *k.* The PLAY score (0-2) is the mean of all 25 practices in the PLAY instrument.

As in equation 1, W_{ijk} is a vector of student characteristics and includes a student's age and gender. X_{jk} is a vector of teacher characteristics including gender, teaching experience in years, and their highest level of education. Finally, Z_k is a vector of school controls including rurality (dummy taking the value of 1 if school is in a rural area), number of students enrolled, average teacher absence rate, infrastructure index, and a school-level wealth index. The infrastructure index measures if a school has access to drinking water, functioning toilet, internet, electricity in the classroom, and how accessible the school is to students with disabilities.

Summary and interpretation

Table 10 shows the coefficients on PLAY and the other indicators of teachers' skill and knowledge, with the different measures of student learning as the dependent variable. As with the Teach regression results in the previous section, the first of each pair of columns shows the correlation with only the country fixed effects, while the second column shows the relationship with full controls. Unlike with the Teach regressions, the sample for these regressions is limited to data from the 3 countries for which we have PLAY data: Ethiopia, Peru, and Sierra Leone.

The results show that a teacher's PLAY score predicts her students' learning both in literacy and overall student knowledge (Table 10, Panel A), and the relationship is both quantitatively meaningful and robustly statistically significant. Specifically, a 1-point increase in PLAY score is associated with an increase of 13 points (out of 100) on the literacy assessment, even with full controls included, and the coefficient is highly significant. For the overall student score, the associated increase is more than 9 points (with full controls), also highly significant. For mathematics scores, the predicted effect is more than 6 points without controls, but it becomes smaller and insignificant when all controls are included.

For comparison with Teach, Panel B in Table 10 shows the coefficients on Teach when the sample is limited to the 3 PLAY countries. The PLAY and Teach scores are highly correlated, at least in Peru and Ethiopia (see Figure 2), so we would expect results to be similar. Teach scores do retain some predictive power for this limited set of countries, but the coefficients on them are somewhat smaller than for the full sample of countries, and significance disappears when the full set of controls are added across literacy, math, and overall student knowledge.

Because the PLAY and Teach indicators use different scales, we examine how a 1-SD increase in each indicator relates to student learning. For these 3 countries, the standard deviation across teachers is 0.25 for PLAY and 0.7 for Teach. When considering the regressions with full controls (Columns 2 and 4 in Table 10), a 1-SD increase in PLAY is associated with a 2.3-point gain in overall scores and a 3.5-point gain in literacy scores, while a 1-SD increase in Teach is associated with gains of around 1 point on both overall and literacy scores, though these are not significant.

When both the PLAY and Teach scores are entered as independent variables alongside content knowledge (see Panel C, Table 10), we find that PLAY's focus on student engagement maintains a strong association with learning outcomes for overall test scores and literacy scores, highlighting the importance of capturing student engagement in classroom observations.

Restricting the sample by the subject taught attenuates some of the results. For example, results from restricting the teacher sample in the literacy regression to teachers who teach language or both math and language show that a 1-point increase in PLAY predicts an increase of 11.7 points in literacy, somewhat lower than the result of 13.8 seen for the full sample (Tables 9 and 10). Controlling for content knowledge, the estimate on literacy diminishes from 13.9 points to 10.4 points. Nevertheless, PLAY retains its significance as a predictor of both literacy and overall scores. For Teach, literacy results become stronger and more significant with the restricted sample (Panel B), but diminish when PLAY and content knowledge are controlled for (Panel C).

The relationship between support for student engagement (as measured by PLAY) and learning outcomes shows interesting patterns across the three countries. When we look at the (binned) correlations between PLAY and overall scores, the relationship for the 3 countries taken together is clear graphically (Figure 2, left graph). When we examine the relationship by country (see Appendix C, Table C7), this relationship is strongest in Ethiopia (Figure 2, right graph): even with full controls and teacher content knowledge added, a 1-point improvement in the PLAY score predicts increases of more than 20 points in overall test scores (significant at 0.01 level) (see Annex C, Table C7, Panel C). While the correlations for the other countries are less obvious in the figure, multivariate regressions show a modestly significant correlation in Sierra Leone (8 points of the overall score increase for each 1-point PLAY increase at the 10% significance level), and a suggestive correlation in Peru (insignificant, but with a similar magnitude of 5 points, which diminishes completely when content knowledge is controlled for in Panel C).

This relationship between pedagogical practices and student learning, though strong overall, is correlational. As it codes the interaction between teachers and students and the practice of students, the PLAY tool could simply be giving higher scores to teachers who happen to have been assigned students who already had stronger learning foundations or were already more oriented toward learning—which could make it easier for the teacher to use practices that support engagement. In that case, the correlation between better PLAY scores and learning could have an element of reverse causation. Because the GEPD data is cross-sectional, it is not possible to eliminate this possibility. The same issue is also applicable to Teach, where the Teach tool could be giving higher scores to teachers who happen to have students with stronger learning foundations, making it easier for the teacher to implement practices highlighted in the Teach tool. To investigate the relationship further, it will be necessary to use experimental or quasi-experimental methods that will allow causal inference. For example, multiple applications of GEPD data collection following the same teachers and students would allow measurement of teacher value-added over time. In the meantime, these results are at least suggestive that PLAY is a good measure of pedagogical practices that can spur foundational learning.

Table 10: Regression Results for Core Model

(a) Full Sample

	(1) Student knowledge	(2) Student knowledge	(3) Literacy student knowledge	(4) Literacy student knowledge	(5) Math student knowledge	(6) Math student knowledge
Panel A: Teacher PLAY score						
PLAY score	10.5***	9.14***	14.4***	13.8***	6.65**	4.46
	(2.81)	(2.59)	(3.29)	(3.07)	(2.85)	(2.74)
Observations	11818	9283	11818	9283	11818	9283
R^2	0.205	0.284	0.168	0.250	0.161	0.215
Panel B: Teacher pedagogical skill						
Teach score	3.02**	1.42	3.30**	1.48	2.74*	1.36
	(1.34)	(1.41)	(1.50)	(1.60)	(1.42)	(1.51)
Observations	11818	9283	11818	9283	11818	9283
R^2	0.199	0.277	0.157	0.237	0.160	0.214
Panel C: Pedagogical skill and content knowledge						
PLAY score	9.73***	8.88***	14.8***	13.9***	4.66	3.90
	(3.21)	(2.76)	(3.78)	(3.31)	(3.17)	(2.86)
Teach score	1.71	0.37	1.40	-0.28	2.02	1.01
	(1.51)	(1.53)	(1.71)	(1.72)	(1.57)	(1.65)
Content knowledge	0.074**	0.025	0.10***	0.048	0.044	0.0021
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)
Observations	9993	9007	9993	9007	9993	9007
R^2	0.201	0.275	0.171	0.245	0.155	0.205
Panel D: PLAY skill sub-scores						
Support for exploration	3.29*	4.33**	5.69***	6.38***	0.88	2.29
	(1.83)	(1.78)	(2.11)	(2.04)	(1.99)	(2.07)
Support for agency	-0.59	0.63	-4.62	-3.52	3.44	4.77
	(2.99)	(3.07)	(3.42)	(3.63)	(3.24)	(3.30)
Support for connection	6.51***	3.18	8.18***	5.26*	4.84*	1.09
	(2.48)	(2.62)	(2.80)	(2.90)	(2.61)	(2.83)
Support for emotional climate	0.42	-0.10	2.93	3.26	-2.09	-3.47
	(2.41)	(2.26)	(2.86)	(2.72)	(2.48)	(2.41)
Observations	11818	9283	11818	9283	11818	9283
R^2	0.208	0.286	0.173	0.253	0.164	0.218
Panel E: Teach skill sub-scores						
Classroom culture	0.25	-1.57	0.088	-2.17	0.40	-0.97
	(1.42)	(1.46)	(1.70)	(1.81)	(1.42)	(1.42)
Instruction	0.38	1.23	0.47	0.97	0.28	1.49
	(1.49)	(1.41)	(1.69)	(1.62)	(1.57)	(1.53)
Socio emotional skills	2.54*	1.51	2.88	2.51	2.20	0.51
	(1.50)	(1.55)	(1.78)	(1.82)	(1.55)	(1.69)
Observations	11818	9283	11818	9283	11818	9283
R^2	0.200	0.279	0.158	0.240	0.161	0.215
School characteristics	No	Yes	No	Yes	No	Yes
Teacher characteristics	No	Yes	No	Yes	No	Yes
Student characteristics	No	Yes	No	Yes	No	Yes
Extended controls	No	No	No	No	No	No
Survey time	No	Yes	No	Yes	No	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

This table presents point estimates from regressing student learning (student knowledge score) on dimensions of teaching quality using our core model. Columns 1 and 2 refer to overall student knowledge, columns 3 and 4 refer to student knowledge in literacy and columns 5 and 6 refer to student knowledge in mathematics. In panel A, teacher pedagogical skill is measured by the PLAY score (0-2), an average of all items in the PLAY instrument. In panel B, teacher pedagogical skill is measured by the TEACH score (1-5), an average of all items in the TEACH instrument. In panel C, teacher content knowledge is added which measures the percentage of correct questions on the teacher content knowledge test (language for language teachers and mathematics for mathematics teachers), ranging from 0 to 100. Panel D includes PLAY skills sub-scores, while panel E includes pedagogical skills sub-scores. In columns 2,4 and 6, the model controls for a vector of student, teacher, and school characteristics. Student controls include age and gender. Teacher controls include gender, teaching experience in years, and teacher's level of education. School controls include rurality, number of students enrolled, average teacher absence rate, infrastructure index, and a school-level wealth index. The infrastructure index measures whether a school has access to drinking water, functioning toilet, functioning internet, and functioning electricity in the classroom, and whether the school is accessible to students with disabilities. These models also control for survey time, the number of months from the start of the academic year to implementation of survey instrument. All models include country fixed effects. Standard errors clustered by school in parentheses. * p< 0.1, ** p< 0.05, *** p< 0.01

(b) Restricted Sample Based on Teacher Subject(s) Taught

	(1) Student knowledge	(2) Student knowledge	(3) Literacy student knowledge	(4) Literacy student knowledge	(5) Math student knowledge	(6) Math student knowledge
Panel A: Teacher PLAY score						
PLAY score	6.37*	6.52**	10.7***	11.7***	7.00**	3.55
	(3.47)	(2.90)	(3.74)	(3.30)	(3.09)	(2.90)
Observations	5830	4795	8181	6848	8802	7180
R^2	0.343	0.423	0.214	0.291	0.197	0.249
Panel B: Teacher pedagogical skill						
Teach score	5.72***	2.32	5.59***	3.64**	2.64*	-0.34
	(1.43)	(1.48)	(1.57)	(1.64)	(1.57)	(1.58)
Observations	5830	4795	8181	6848	8802	7180
R^2	0.350	0.420	0.215	0.285	0.195	0.248
Panel C: Pedagogical skill and content knowledge						
PLAY score	2.97	5.54*	9.89**	10.4***	4.24	3.94
	(3.92)	(3.01)	(4.24)	(3.57)	(3.36)	(2.96)
Teach score	3.34**	0.94	3.30*	2.31	1.11	-0.85
	(1.65)	(1.60)	(1.77)	(1.73)	(1.75)	(1.72)
Content knowledge	0.13***	0.064*	0.14***	0.067*	0.050	-0.0046
	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
Observations	4964	4630	7162	6623	7745	6964
R^2	0.354	0.423	0.222	0.290	0.183	0.240
Panel D: PLAY skill sub-scores						
Support for exploration	1.75	3.43*	4.08*	4.32**	0.64	2.18
	(2.15)	(2.06)	(2.33)	(2.17)	(2.13)	(2.13)
Support for agency	-1.43	-0.080	-5.34	-2.65	3.55	3.30
	(3.60)	(3.58)	(3.72)	(3.80)	(3.64)	(3.72)
Support for connection	8.70***	3.83	11.0***	7.02**	4.27	0.89
	(2.82)	(2.93)	(2.97)	(2.97)	(2.86)	(3.06)
Support for emotional climate	-4.75	-2.07	-1.40	1.48	-1.03	-2.88
	(2.93)	(2.92)	(3.28)	(3.17)	(2.60)	(2.53)
Observations	5830	4795	8181	6848	8802	7180
R^2	0.354	0.426	0.224	0.294	0.199	0.251
Panel E: Teach skill sub-scores						
Classroom culture	2.79	-0.47	0.76	-1.75	1.04	-0.70
	(1.85)	(1.51)	(2.00)	(2.01)	(1.71)	(1.58)
Instruction	0.69	2.36*	2.61	3.79**	-0.64	0.44
	(1.67)	(1.43)	(1.79)	(1.72)	(1.67)	(1.58)
Socio emotional skills	2.66	-0.20	1.95	0.57	2.65	-0.23
	(1.82)	(1.83)	(1.90)	(1.94)	(1.70)	(1.75)
Observations	5830	4795	8181	6848	8802	7180
R^2	0.351	0.421	0.215	0.288	0.196	0.249
School characteristics	No	Yes	No	Yes	No	Yes
Teacher characteristics	No	Yes	No	Yes	No	Yes
Student characteristics	No	Yes	No	Yes	No	Yes
Extended controls	No	No	No	No	No	No
Survey time	No	Yes	No	Yes	No	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

This table presents point estimates from regressing student learning (student knowledge score) on dimensions of teaching quality using our core model. In this table, we restrict teachers depending on the subject taught: columns 1 and 2 include only teachers who teach both language and math, columns 3 and 4 include teachers who teach either both language and math or language only, and columns 5 and 6 include teachers who teach either language and math or math only. Columns 1 and 2 refer to overall student knowledge, columns 3 and 4 refer to student knowledge in literacy and columns 5 and 6 refer to student knowledge in mathematics. In panel A, teacher pedagogical skill is measured by the Play score (0-2), an average of all items in the Play instrument. In panel B, teacher pedagogical skill is measured by the TEACH score (1-5), an average of all items in the TEACH instrument. In panel C, teacher content knowledge is added which measures the percentage of correct questions on the teacher content knowledge test (language for language teachers and mathematics for mathematics teachers), ranging from 0 to 100. Panel D includes PLAY skills sub-scores, while panel E includes pedagogical skills sub-scores. In columns 2,4 and 6, the model controls for a vector of student, teacher, and school characteristics. Student controls include age and gender. Teacher controls include gender, teaching experience in years, and teacher's level of education. School controls include rurality, number of

students enrolled, average teacher absence rate, infrastructure index, and a school-level wealth index. The infrastructure index measures whether a school has access to drinking water, functioning toilet, functioning internet, and functioning electricity in the classroom, and whether the school is accessible to students with disabilities. These models also control for survey time, the number of months from the start of the academic year to implementation of survey instrument. All models include country fixed effects. Standard errors clustered by school in parentheses. * p< 0.1, ** p< 0.05, *** p< 0.01

It is also useful to explore the explanatory power of each of the 4 main subskills measured by PLAY, especially because PLAY is a recently developed tool. Panel D of Table 10 shows the coefficients on those subskills, from a version of the Panel A regressions in which the four subskills are included as regressors, in place of the composite PLAY score. Like the composite score, the individual subskill scores are more predictive of student scores on literacy than on numeracy. When we include all controls, the subskills for *support for exploration* and *support for connection* are both strongly predictive of better literacy scores. By contrast, *support for agency* and *support for emotional climate* are not significant.

Teacher Support Matters for Improved Pedagogical Practices

The results above suggest that teaching quality, as proxied by the indicators of teacher pedagogical skill and teacher content knowledge, is generally predictive of student learning in low- and middle-income countries. This highlights the importance of understanding how countries can improve teachers' pedagogical skill and content knowledge. In this section, we explore this question by using the rich set of GEPD variables to identify which factors at various levels best predict each of the 2 outcomes.

How we estimate the drivers of pedagogical quality and teacher content knowledge

The models used to estimate the two outcomes use a different sample. To estimate the correlates of pedagogical skill, we restrict our sample to the 2,686 teachers selected for the classroom observation module. When our outcome is teacher content knowledge, our sample is 11,755 teachers who were assessed in either literacy or numeracy. Our model specification is as follows:

$$Y_{ijks} = \beta_1(A_{jk}) + \beta_2(B_{jk}) + \beta_3(C_{ijk}) + \beta_4(D_{ijk}) + \beta_5(E_{ijk}) + \beta_6(F_{ijk}) + \beta_7(G_{ijk}) + \beta_8(H_{jk}) + \varepsilon \qquad (3)$$

where Y_{ijks} is teaching quality of teacher *i* in school *j* in country *k,* and teaching quality is denoted by *s = {Teach score, content knowledge score}*.

For the core model, we restrict the explanatory variables to: A_{jk}, a vector of teacher characteristics including gender, teaching experience in years, and their highest level of education; and B_{jk} is a vector of school controls including rurality (dummy taking the value of 1 if school is in a rural area), number of students enrolled, average teacher absence rate, and infrastructure index.

The full model further controls for the following groups of indicators: recruitment and training (C_{ijk}), instructional leadership (D_{ijk}), teacher evaluation (E_{ijk}), incentives (F_{ijk}), intrinsic motivation (G_{ijk}), and school-level factors (H_{jk}). Details of these groups of indicators may be found in Annex B.

In Tables 11, 12, and 13, we show the pooled regression results across all countries for the correlates of teacher pedagogical skill and content knowledge, respectively.

Factors predicting better pedagogy (as measured by Teach)

First, we explore the factors, at both the teacher and school level, that predict which teachers display more effective pedagogical skills in the classroom, as measured by Teach (see Table 11).

Some basic *demographic factors*—teacher education and gender—predict a higher Teach score:

Teachers with more *education* show somewhat better pedagogical quality. Those with a bachelor's degree score higher than those with less education, and the advantage for those with master's or doctoral degrees is about twice as large. That said, the difference is small: even those with postgraduate degrees score only 0.13 points (Table 11, Column 7) higher on a 1-to-5 scale than teachers who completed at most high school, while the teachers with only bachelor's degrees—the most prevalent education level in our sample—score less than 0.1 points higher (Table 11, Column 7). *Female* teachers also score slightly higher, with a statistically robust but small advantage of about 0.06 points (Table 11, Column 7). This effect is driven by lower-middle income countries, where female teachers score 0.1 points higher than their male counterparts (see Table C8 of Appendix C). *School location and environment* also correlate with pedagogy.

- Teachers in *rural* schools score slightly lower on pedagogy than those in *urban* schools (once other demographic and school-environment variables are controlled for), and those at schools with better *infrastructure* also score somewhat better. However, neither effect is large and both appear only in the lower-middle income group.
- Teachers at *larger schools* score significantly higher, although the effect is very small.

Beyond these demographic factors, a number of *pedagogical support variables* are also correlated with higher Teach scores, even when we include an extensive set of control variables. Three of these variables reflect the support that the system can provide for better teaching:

- *Induction/mentorship program:* Participation in an induction or mentorship program is generally positively correlated with a teacher's pedagogical skill (at 5% significance level), except in Column 7 in Table 11 with the full set of controls (where the effect is likely absorbed by the added teacher support and instructional leadership school-level variables). This correlation is consistent with an extensive literature that argues that, especially where pre-service education has not prepared a teacher well, it is essential for experienced teachers or coaches to help a new teacher learn to handle the challenges of actual classroom teaching.
- *Feedback on lesson plans:* Discussing lesson plans also predicts higher Teach scores. Interestingly, merely having a lesson plan does not; instead, it is the process of having discussed and gotten feedback on them that appears to help. This is also consistent with the idea that feedback can help teachers to improve.

The coefficient size on each of these variables is roughly comparable to those on the demographic variables such as gender and rurality, at 0.05 and 0.07 points (Table 11, Column 6). Together, they add up to a substantial impact of 0.12 points on the 1-5 scale. Given that the mean Teach score is 2.7, this represents a 4-percent improvement in scores. Expressed another way, because the standard deviation of the Teach score across countries is 0.7, the predicted effect of better practical pedagogical support is 0.17 of a standard deviation—a nontrivial impact, if the correlation reflects a causal relationship. Another variable representing practical support for classroom practice—the share of in-service training received in the classroom—is not statistically significant, but it has a robustly positive and large coefficient across specifications.

A third group of variables somewhat correlated with Teach scores relates to the *career- and pay-related signals* that the system sends about who and what it values, but these variables have inconsistent predicted effects and are less significant. The *financial consequences to positive evaluation* has a negative coefficient; this is contrary to the expected positive relationship, but the coefficient is only marginally significant in most specifications.

Compensation predictability is also somewhat related to teaching quality, in that teachers whose salary was delayed during the last academic year had lower Teach scores; but again, these coefficients are only marginally significant and become insignificant once school-level de facto policy variables are added.

Finally, two *school-level de facto policy variables*—the quality of *teaching evaluation* and *instructional leadership*—are also significant predictors of higher Teach scores. A higher teaching evaluation score means that the school has an evaluation system in place, more teachers have been evaluated in the past year, and teachers report that the evaluation has consequences. The instructional leadership score reflects whether teachers at the school report that frequent classroom observations were followed by substantial discussions and actionable feedback, and whether teachers had lesson plans that they discussed with another person.

Factors predicting better pedagogy (as measured by PLAY)

Table 12 explores what factors predict PLAY scores, using the same set of regressors. Because PLAY measures teachers' ability to support children's engagement, including through practices that are not always prioritized in teacher pre-service and in-service education, it is not obvious *a priori* that the same factors will predict better pedagogy when pedagogical quality is measured via PLAY. And indeed, the results for PLAY appear to differ substantially from those for Teach. Fewer regressors are statistically significant, and the overall explanatory power is slightly lower. One reason for this is likely that the sample size is smaller: the sample for the PLAY regressions, based on data for 3 three of the 13 GEPD systems, is less than a quarter as large as the sample for the Teach regressions. But there are meaningful differences as well. Teachers' education levels are not significant predictors of PLAY scores, whereas they remain statistically and quantitatively significant predictors of Teach scores even when we restrict the sample to the three PLAY countries (see Table C10 of Appendix C).

Two measures of *pedagogical support* predict better pedagogical practice as measured by PLAY:

- Teachers who have gotten *feedback on lesson plans* have higher PLAY scores. As in the case of Teach, merely having lesson plans does not have any predicted effect. However, as with Teach, the coefficient is not statistically significant with the full set of controls in Column 7, likely because the frequency of feedback on lesson plans is subsumed under the *instructional leadership* variable.
- The composite *de facto* policy variable measuring quality of *instructional leadership* at the school also predicts higher PLAY scores (Table 12, Column 7).

Additionally, for PLAY as for Teach, higher *de facto* policy scores for *teacher evaluation* predict better pedagogical practice.

Finally, teachers who report that they were motivated to go into teaching because of the *steady career path* have lower scores. While the coefficient is only marginally significant, this finding could suggest that teachers who are less motivated are less likely to ensure that children are actively engaged in learning (which PLAY measures).

Factors predicting higher levels of teacher content knowledge

Table 13 shows the correlates of teachers' content knowledge.

First, on the *demographic* factors:

- *More educated* teachers score higher on the content knowledge assessment, as they do on the Teach observation. Teachers with at least a bachelor's degree score more than 4 points higher (out of 100) compared to those with only a high-school diploma or less. Surprisingly, those with a graduate degree

have a similar score compared to those with only a bachelor's degree. Given that the average content knowledge score across all teachers is only 54, this similarity is not likely due to ceiling effects. The predicted effect of education is especially pronounced in upper-middle-income countries, where teachers with a bachelor's degree score 9.7 points higher on content knowledge than do teachers with high school diploma or less, while teachers with a master's or doctoral degree score 12 points higher (see Table C9 of Annex C).

- *Male* teachers do slightly better on content knowledge than do female teachers, scoring 1.3 points (out of 100) higher, even when we include all control variables (Table 13, Column 7). This effect is driven mostly by the low-income countries in the sample; however, in the lower-middle-income countries, female teachers score higher than their male peers (Table C9, Column 1).

Second, *school location, environment,* and *size* variables are significant predictors of teachers' content knowledge:

- Teachers in *rural* schools score lower than those in urban schools (once we control for factors like teacher education, school size, and infrastructure quality), and *infrastructure* is positively correlated with scores.
- In contrast to our findings for Teach, teachers in *smaller* schools score somewhat better on content knowledge.

Third, unlike with Teach, *pedagogical support and preparation* variables do not show a clear and easily interpretable relationship with content knowledge. For example:

- *In-service training* has ambiguous and offsetting effects: having participated in recent training is associated with *lower* content knowledge scores, while the number of days of training is positively associated with content knowledge, though the magnitude is small. Teachers whose training included a significant in-classroom component (over 50%) score 4 points higher in content knowledge than other teachers, with an even larger effect in low-income countries (Table 13, Column 7 and Table C9, Column 1).
- *Classroom observation* is not associated with higher content knowledge. Teachers whose classroom had been observed scored significantly lower, while those who had discussed the observation scored somewhat higher but that result is not significant.
- Teachers who had *lesson plans* for the previous week earned higher scores, suggesting an association between knowledge and preparation (or support, depending on the system). Lesson plans can be an important tool to help teachers deliver lessons in an effective manner by providing a structured framework to support teachers on both pedagogy and content. The effect of discussing lesson plans is positive but no longer statistically significant.

Fourth, teachers *motivated* to teach—whether because they enjoy it, see teaching as a steady career path, or believe teaching benefits others—score substantially higher than those who did not report any reason for becoming a teacher.

In short, there is less of a clear explanatory story for content knowledge than for pedagogical skill, at least from a policy perspective. While teachers with at least a bachelor's degree score somewhat higher, as do those who report some source of motivation, there is little indication that variables related to support, evaluation, and monitoring are correlated with content knowledge in any robust and consistent way.

Table 11: Regression Results – Teacher Pedagogical Skills

	(1) Teach score	(2) Teach score	(3) Teach score	(4) Teach score	(5) Teach score	(6) Teach score	(7) Teach score
Female teacher	0.047*	0.051**	0.049**	0.051**	0.052**	0.056**	0.063**
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Teaching experience (in	-0.0012	-0.0012	-0.0012	-0.0012	-0.0013	-0.0015	-0.0013

years)							
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Highest education: Bachelors	0.091**	0.087**	0.086**	0.085**	0.080**	0.080**	0.073**
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Highest education: master or doctoral	0.18***	0.16***	0.16***	0.16***	0.15***	0.15***	0.13**
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Highest education: Other	0.018	0.028	0.025	0.026	0.024	0.024	0.015
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Rurality	-0.064**	-0.069**	-0.063**	-0.063**	-0.059**	-0.056**	-0.058**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Infrastructure index	0.035***	0.031***	0.029**	0.028**	0.027**	0.027**	0.022*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Students enrolled	0.000085***	0.000085***	0.000087***	0.000086***	0.000088***	0.000091***	0.000091***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Average teacher absence rate (school level)	-0.000046	0.000019	0.00012	0.00012	0.00012	0.00013	0.00037
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Recruitment score (0-9)		-0.0057	-0.0075	-0.0062	-0.0064	-0.0076	-0.011
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Probationary period for new teachers		-0.0017	-0.0097	-0.0087	-0.014	-0.013	-0.042
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Participated in induction/mentorship program		0.059**	0.050**	0.051**	0.055**	0.051**	0.036
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Teaching practicum under pre-service training		0.049	0.039	0.039	0.031	0.025	0.0067
		(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Hours per day spent teaching during teaching practicum		-0.0044	-0.0044	-0.0048	-0.0039	-0.0034	-0.0035
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Participated in in-service training		0.024	0.016	0.015	0.014	0.016	0.0014
		(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Days in-service training lasted		-0.0010	-0.0012	-0.0012	-0.0011	-0.0012	-0.0010
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Less than 50% in service training in classroom		0.0022	-0.012	-0.0088	-0.015	-0.020	-0.024
		(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
Over 50% in service training in classroom		0.092*	0.084	0.088	0.093*	0.10*	0.092*
		(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)
Classroom has been observed			0.036	0.040	0.044	0.038	0.014
			(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Months since last classroom observation			0.00019	0.00015	0.00012	0.00017	0.0027
			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Discussed results of observation			0.090	0.088	0.086	0.086	0.037
			(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Observer provided feedback			-0.058	-0.055	-0.054	-0.049	-0.049
			(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Had lesson plans for last week			0.0043	0.0060	0.0052	0.0047	-0.022
			(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Discussed lesson plans			0.071***	0.071***	0.069**	0.071***	0.057*
			(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Evaluated on pedagogical skill and content knowledge				-0.011	-0.012	-0.012	-0.047*
				(0.03)	(0.02)	(0.02)	(0.03)
Financial consequences to negative evaluation				-0.0030	-0.0066	0.00064	-0.0017
				(0.03)	(0.03)	(0.03)	(0.03)
Financial consequences to positive evaluation				-0.040	-0.051*	-0.053*	-0.056**
				(0.03)	(0.03)	(0.03)	(0.03)
Meritocracy for advancement					0.043*	0.039	0.0074
					(0.02)	(0.02)	(0.03)
Received bonus during last AY					0.062	0.065	0.058

					(0.04)	(0.04)	(0.05)
Salary was delayed during last AY					-0.056*	-0.059*	-0.052
					(0.03)	(0.03)	(0.03)
Motivation - like teaching						-0.18	-0.17
						(0.12)	(0.12)
Motivation - steady career path						-0.18	-0.14
						(0.13)	(0.13)
Motivation - teaching benefits others						-0.093	-0.083
						(0.12)	(0.12)
Intrinsic motivation (school level)							0.055*
							(0.03)
Teacher monitoring (school level)							-0.030
							(0.02)
Teaching evaluation (school level)							0.039**
							(0.02)
Teacher support (school level)							0.026
							(0.02)
Teacher attraction (school level)							0.051
							(0.04)
Instructional leadership (school level)							0.038**
							(0.02)
Constant	2.51***	2.49***	2.43***	2.44***	2.43***	2.60***	2.12***
	(0.05)	(0.06)	(0.06)	(0.07)	(0.07)	(0.14)	(0.22)
Observations	2137	2135	2135	2134	2134	2134	2092
R^2	0.252	0.261	0.268	0.268	0.271	0.275	0.281
School characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Recruitment and Training	No	Yes	Yes	Yes	Yes	Yes	Yes
Instructional leadership	No	No	Yes	Yes	Yes	Yes	Yes
Evaluation	No	No	No	Yes	Yes	Yes	Yes
Incentives	No	No	No	No	Yes	Yes	Yes
Intrinsic motivation	No	No	No	No	No	Yes	Yes
School level indicators	No	No	No	No	No	No	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

For the highest level of education completed, the omitted category is high school and below. For the percent of in-service training done in classroom, the omitted category is that no in-service training was done in the classroom. A recruitment score sums up the number of questions a teacher responded yes to under the recruitment category (Table 8). An infrastructure index measures if a school has access to drinking water, functioning toilet, internet, electricity in the classroom and how accessible the school is to students with disabilities.* p<0.01 ** p<0.05 *** p<0.10. Standard errors clustered at school level in parentheses.

Table 12: Regression Results – Teacher PLAY Score

	(1) PLAY score	(2) PLAY score	(3) PLAY score	(4) PLAY score	(5) PLAY score	(6) PLAY score	(7) PLAY score
Female teacher	-0.016	-0.015	-0.016	-0.015	-0.017	-0.017	-0.0063
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Teaching experience (in years)	0.0011	0.00066	0.00069	0.00065	0.00065	0.00062	0.0011
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Highest education: Bachelors	0.00038	0.0052	0.0023	0.0032	0.012	0.010	-0.0066
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Highest education: master or doctoral	0.0073	0.012	0.00032	0.0021	0.0053	0.0042	-0.015
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)
Highest education: Other	-0.0058	0.0023	-0.00020	0.0017	0.0073	0.0081	-0.015
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Rurality	-0.037	-0.035	-0.035	-0.036	-0.032	-0.031	-0.030
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Infrastructure index	0.014	0.018*	0.015	0.014	0.014	0.013	0.0086
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)

Students enrolled	-0.000023	-0.000030	-0.000028	-0.000027	-0.000025	-0.000023	-0.000026
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Average teacher absence rate (school level)	-0.00046	-0.00037	-0.00033	-0.00029	-0.00026	-0.00031	-0.00025
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Recruitment score (0-9)		-0.0067	-0.0086	-0.0084	-0.0075	-0.0069	-0.010
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Probationary period for new teachers		-0.014	-0.015	-0.014	-0.0092	-0.0097	-0.011
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Participated in induction/mentorship program		0.022	0.023	0.023	0.024	0.022	0.021
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Teaching practicum under pre-service training		0.035	0.034	0.033	0.031	0.032	0.027
		(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Hours per day spent teaching during teaching practicum		-0.010	-0.011	-0.011	-0.011	-0.011	-0.0090
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Participated in in-service training		0.030	0.033	0.034	0.036	0.032	0.027
		(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Days in-service training lasted		0.00057	0.00045	0.00046	0.00057	0.00057	0.00080
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Less than 50 in service training in classroom		-0.0097	-0.018	-0.019	-0.020	-0.018	-0.030
		(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Over 50 in service training in classroom		-0.021	-0.016	-0.018	-0.014	-0.013	-0.034
		(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)
Classroom has been observed			0.0059	0.0050	0.0055	0.0095	-0.011
			(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Months since last classroom observation			0.0010	0.0010	0.0011	0.00070	0.0020
			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Discussed results of observation			0.0052	0.0044	-0.0029	-0.0049	-0.060
			(0.06)	(0.06)	(0.06)	(0.05)	(0.05)
Observer provided feedback			-0.029	-0.029	-0.023	-0.023	-0.031
			(0.05)	(0.05)	(0.05)	(0.05)	(0.04)
Had lesson plans for last week			-0.0031	-0.0026	0.00092	-0.00038	0.0042
			(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Discussed lesson plans			0.061***	0.060***	0.060***	0.058**	0.031
			(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Evaluated on pedagogical skill and content knowledge				0.0055	0.0023	0.00033	-0.035
				(0.02)	(0.02)	(0.02)	(0.03)
Financial consequences to negative evaluation				0.0066	0.0035	-0.0011	0.0070
				(0.03)	(0.03)	(0.03)	(0.03)
Financial consequences to positive evaluation				-0.014	-0.0099	-0.0068	-0.0095
				(0.02)	(0.02)	(0.02)	(0.02)
Meritocracy for advancement					-0.022	-0.022	0.0032
					(0.02)	(0.02)	(0.03)
Received bonus during last AY					0.026	0.029	0.046
					(0.04)	(0.04)	(0.04)
Salary was delayed during last AY					-0.039	-0.039	-0.048*
					(0.03)	(0.03)	(0.03)
Motivation - like teaching						-0.045	-0.055
						(0.04)	(0.04)
Motivation - steady career path						-0.11*	-0.097*
						(0.06)	(0.05)
Motivation - teaching benefits others						-0.045	-0.048
						(0.05)	(0.05)
Intrinsic motivation (school							0.026

level)							
							(0.03)
Teacher monitoring (school level)							-0.0092
							(0.02)
Teaching evaluation (school level)							0.044*
							(0.02)
Teacher support (school level)							0.00032
							(0.02)
Teacher attraction (school level)							-0.032
							(0.03)
Instructional leadership (school level)							0.056***
							(0.02)
Constant	0.37***	0.37***	0.35***	0.35***	0.36***	0.41***	0.18
	(0.05)	(0.05)	(0.07)	(0.07)	(0.07)	(0.09)	(0.18)
Observations	538	538	538	538	538	538	537
R^2	0.198	0.209	0.220	0.221	0.226	0.231	0.262
School characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Recruitment and Training	No	Yes	Yes	Yes	Yes	Yes	Yes
Instructional leadership	No	No	Yes	Yes	Yes	Yes	Yes
Evaluation	No	No	No	Yes	Yes	Yes	Yes
Incentives	No	No	No	No	Yes	Yes	Yes
Intrinsic motivation	No	No	No	No	No	Yes	Yes
School level indicators	No	No	No	No	No	No	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

For the highest level of education completed, the omitted category is high school and below. For the percent of in-service training done in classroom, the omitted category is that no in-service training was done in the classroom. A recruitment score sums up the number of questions a teacher responded yes to under the recruitment category (Table 87). An infrastructure index measures if a school has access to drinking water, functioning toilet, internet, electricity in the classroom and how accessible the school is to students with disabilities.* p<0.01 ** p<0.05 *** p<0.10. Standard errors clustered at school level in parentheses.

Table 13: Regression Results – Teacher Content Knowledge

	(1) Content knowledge	(2) Content knowledge	(3) Content knowledge	(4) Content knowledge	(5) Content knowledge	(6) Content knowledge	(7) Content knowledge
Female teacher	-1.31**	-1.41***	-1.44***	-1.44***	-1.42***	-1.42***	-1.32**
	(0.53)	(0.53)	(0.53)	(0.53)	(0.53)	(0.53)	(0.54)
Teaching experience (in years)	0.011	0.0055	0.0060	0.0087	0.0067	0.0081	0.015
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Highest education: Bachelors	4.69***	4.73***	4.77***	4.77***	4.77***	4.82***	4.77***
	(0.90)	(0.90)	(0.90)	(0.89)	(0.89)	(0.89)	(0.90)
Highest education: master or doctoral	4.61***	4.64***	4.60***	4.63***	4.61***	4.65***	4.75***
	(1.02)	(1.02)	(1.02)	(1.01)	(1.01)	(1.01)	(1.02)
Highest education: Other	2.89***	3.02***	3.06***	2.98***	3.00***	3.03***	2.86***
	(0.97)	(0.96)	(0.95)	(0.95)	(0.95)	(0.95)	(0.97)
Rurality	-2.16***	-2.21***	-2.21***	-2.17***	-2.13***	-2.12***	-2.12***
	(0.69)	(0.69)	(0.69)	(0.69)	(0.69)	(0.69)	(0.69)
Infrastructure index	1.10***	1.05***	1.03***	1.06***	1.05***	1.06***	0.99***
	(0.32)	(0.32)	(0.32)	(0.32)	(0.32)	(0.32)	(0.32)
Students enrolled	-0.0018**	-0.0017**	-0.0018**	-0.0017**	-0.0017**	-0.0017**	-0.0019***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Average teacher absence rate (school level)	0.0024	0.0044	0.0045	0.0038	0.0039	0.0048	0.00032
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Recruitment score (0-9)		0.0087	0.015	0.0088	0.0070	0.0043	-0.040
		(0.17)	(0.17)	(0.17)	(0.17)	(0.17)	(0.17)
Probationary period for new teachers		-0.25	-0.30	-0.36	-0.38	-0.35	-0.92
		(0.54)	(0.53)	(0.53)	(0.53)	(0.53)	(0.59)
Participated in induction/mentorship		-0.78	-0.76	-0.73	-0.70	-0.70	-0.82

program						
	(0.52)	(0.51)	(0.51)	(0.51)	(0.51)	(0.53)
Teaching practicum under pre-service training	1.12	1.15	1.22	1.15	1.12	0.65
	(0.87)	(0.87)	(0.87)	(0.87)	(0.87)	(0.89)
Hours per day spent teaching during teaching practicum	-0.28	-0.28	-0.29	-0.28	-0.28	-0.20
	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)	(0.18)
Participated in in-service training	-2.16***	-2.05***	-2.01***	-2.00***	-2.02***	-2.00***
	(0.66)	(0.66)	(0.66)	(0.66)	(0.66)	(0.67)
Days in-service training lasted	0.096***	0.093***	0.094***	0.095***	0.094***	0.094***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Less than 50% in service training in classroom	1.68*	1.52	1.47	1.44	1.41	1.41
	(1.01)	(1.01)	(1.01)	(1.01)	(1.01)	(1.01)
Over 50% in service training in classroom	4.35***	4.20***	4.23***	4.25***	4.29***	4.03***
	(1.14)	(1.14)	(1.14)	(1.14)	(1.14)	(1.14)
Classroom has been observed		-1.71**	-1.62**	-1.60**	-1.62**	-1.70**
		(0.76)	(0.75)	(0.75)	(0.75)	(0.76)
Months since last classroom observation		-0.011	-0.011	-0.010	-0.011	-0.0067
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Discussed results of observation		1.26	1.27	1.23	1.28	1.12
		(1.06)	(1.06)	(1.06)	(1.07)	(1.07)
Observer provided feedback		-0.25	-0.25	-0.21	-0.21	-0.39
		(0.96)	(0.96)	(0.96)	(0.96)	(0.97)
Had lesson plans for last week		1.97***	1.97***	1.95***	1.98***	1.86***
		(0.70)	(0.70)	(0.70)	(0.70)	(0.71)
Discussed lesson plans		0.049	0.068	0.067	0.062	-0.024
		(0.58)	(0.58)	(0.58)	(0.58)	(0.59)
Evaluated on pedagogical skill and content knowledge			-0.41	-0.41	-0.42	-1.10**
			(0.53)	(0.53)	(0.53)	(0.56)
Financial consequences to negative evaluation			0.88	0.88	0.92	0.89
			(0.73)	(0.73)	(0.73)	(0.74)
Financial consequences to positive evaluation			0.61	0.57	0.55	0.39
			(0.58)	(0.59)	(0.59)	(0.60)
Meritocracy for advancement				0.073	0.060	-0.072
				(0.54)	(0.53)	(0.58)
Received bonus during last AY				0.32	0.31	0.38
				(0.87)	(0.87)	(0.87)
Salary was delayed during last AY				-0.82	-0.77	-0.65
				(0.70)	(0.70)	(0.71)
Motivation - like teaching					5.38**	5.45**
					(2.25)	(2.23)
Motivation - steady career path					6.65***	7.17***
					(2.33)	(2.31)
Motivation - teaching benefits others					6.00***	6.04***
					(2.28)	(2.26)
Intrinsic motivation (school level)						1.16
						(0.83)
Teacher monitoring (school level)						-0.043
						(0.57)
Teaching evaluation (school level)						0.90*
						(0.50)
Teacher support (school level)						0.21
						(0.51)
Teacher attraction (school						-0.063

level)							
							(0.87)
Instructional leadership (school level)							0.24
							(0.44)
Constant	53.3***	53.9***	53.0***	52.7***	52.8***	47.1***	40.1***
	(1.41)	(1.52)	(1.61)	(1.59)	(1.62)	(2.65)	(4.82)
Observations	9558	9555	9553	9552	9552	9552	9416
R^2	0.178	0.182	0.184	0.185	0.185	0.186	0.187
School characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Recruitment and Training	No	Yes	Yes	Yes	Yes	Yes	Yes
Instructional leadership	No	No	Yes	Yes	Yes	Yes	Yes
Evaluation	No	No	No	Yes	Yes	Yes	Yes
Incentives	No	No	No	No	Yes	Yes	Yes
Intrinsic motivation	No	No	No	No	No	Yes	Yes
School level indicators	No	No	No	No	No	No	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

For the highest level of education completed, the omitted category is high school and below. For the percent of in-service training done in classroom, the omitted category is that no in-service training was done in the classroom. A recruitment score sums up the number of questions a teacher responded yes to under the recruitment category (Table 8). An infrastructure index measures if a school has access to drinking water, functioning toilet, internet, electricity in the classroom and how accessible the school is to students with disabilities.* p<0.01 ** p<0.05 *** p<0.10. Standard errors clustered at school level in parentheses.

Conclusion

What kinds of teacher pedagogical practices can improve foundational learning—including basic literacy and numeracy skills—of students in primary schools in low- and middle-countries? And are there promising policy levers for improving the pedagogical skills and content knowledge of teachers? As many countries grapple with a foundational learning crisis that has been worsened by pandemic-driven school closures, these questions are of first-order importance.

In this paper, we provide new evidence on these questions using a unique dataset from 13 low- and middle-income-country education systems. The data was collected recently through the World Bank's Global Education Policy Dashboard, using direct observation, student and teacher assessments, and teacher and principal interviews in a representative random sample of schools in each country. These are complemented by data from a policy survey and a survey of civil servants to map out the policy and political climate. Together, the data allow us to draw on an unusually comprehensive set of student, teacher, school, and policy variables to investigate teaching, learning, and their drivers.

First, we map out the *teaching and learning landscape* in these 13 systems:

- *Learning outcomes* of Grade 4 students on a test of foundational literacy and numeracy are quite low in most cases, aligning with what international assessments have highlighted in recent years—learning outcomes are poor across the board. On our assessment—which includes items with a range of difficulty, from roughly Grade 1 up to Grade 4 level, students performed better in literacy (65%) than in mathematics (40%). Many Grade 4 students have mastered the earliest skills, such as addition and word recognition, but struggle with more complex ones like division or comprehension of even short reading passages.
- *Pedagogical skills* of teachers are also weak. The results on teachers' pedagogical practices in classrooms, as measured by the Teach classroom observation tool, show that none of the surveyed education systems had a satisfactory average Teach score (3.5 or higher on a 1-to-5 scale) among its teachers. Only 4 of 13 systems had an average score of 3 or higher. Teachers typically scored relatively high on fostering supportive learning environments. In contrast, teachers struggled to promote social and collaborative skills of students and qualities such a perseverance.

- The results on the prevalence of *teaching practices that support student engagement with learning*, as measured by the PLAY classroom observation tool, show a similar pattern. For example, teachers do better on practices related to support for emotional climate (with items like the use of positive discipline and rewards to support a positive emotional climate, inclusion of children with diverse backgrounds and learning needs, as well as the use of a mode of instruction that is explicitly upbeat) than they do on practices to support the agency of students.
- *Teacher content knowledge* is low, with the average teacher showing mastery of only a little more than half of the Grade 4 content they teach, in both literacy and math. Across the 13 systems, teachers answered 54% of the questions correctly, with a low of 38% in Sierra Leone and a high of 70% in Punjab, Pakistan.

Second, issues with quality of teaching can be exacerbated by a lack of system support for teachers through *training and instructional leadership*. The GEPD data show that this is a widespread issue. For example:

- Although research shows that quality *teaching practicums* during pre-service teacher education are important to improve teaching, they remain rare in these low- and middle-income countries: in 9 of the 13 systems, less than 50% of their teachers having benefited from one.
- *In-service training* often does not reach every teacher: of the 13 systems, only 3 (Peru, Rwanda, and Pakistan's ICT province) had provided in-service training to at least 50% of teachers over the previous 12 months.
- Perhaps more importantly, even when in-service training does take place, it is often not likely to be practical. Although research shows that practical *in-classroom training* is most likely to spur better teaching and learning, 85% of teachers on average across these systems report their last in-service training had no in-classroom component.
- Being observed and getting feedback on pedagogy can help a teacher improve. Encouragingly, 75% of teachers on average reported that their classroom was observed over the past 12 months. However, the follow-up is often weak: only 54% of teachers reported receiving feedback after the classroom observation.

Third, the results show that the Teach and PLAY measures of pedagogical quality do *predict student learning*:

- Students' learning is positively associated with the pedagogical skills of their teachers. This remains true even when we include a wide range of controls at the teacher and school level, including many controls that other datasets do not generally include.
- Pedagogical skills, as measured by Teach, predict students' scores more robustly on the literacy test than for mathematics. Also, the association is much stronger for the middle-income countries in our sample, on average, than for the low-income countries.
- Pedagogical strategies that support student engagement, as measured by PLAY, also appear to predict student learning in the three countries for which data is available. A 1-SD increase in a teacher's PLAY score increases the predicted literacy score by about 3.35 points (or 5%) on average. These results underscore the importance of measuring and supporting aspects related to student engagement with the instructional content.

Finally, the analysis identifies *policy variables* that predict better pedagogical quality:

- Exposure to some *practical pedagogical support interventions* is associated with better pedagogical quality. The interventions include participating in an induction or mentoring program, receiving feedback on lesson plans, and having received recent practical teacher training with a classroom component.

Compared with other interventions, these three are likely to focus more on giving teachers practical skills that will help them teach more effectively in the classroom. A teacher who has benefited from all 3 will score about 8% (or nearly 1/3 of a standard deviation) higher on the Teach scale.

- Consistent with these results, a school-wide summary measure of the *quality of instructional leadership* (derived from teacher interviews) is also a highly significant predictor of higher Teach scores, even with a full set of controls.
- Other variables amenable to policy, such as the degree of perceived *meritocracy* in career advancement, also predict better pedagogy. Higher levels of *education* also are correlated with higher scores on Teach (though not on PLAY).

Given the correlation between pedagogical quality (as measured by Teach and PLAY) and student learning, it is critical to strengthen the support teachers receive inside and outside the classroom. Based on our results, this support could include expanding the percentage of teachers benefiting from high-quality training and practicums, ensuring that training is practical and classroom-based, and providing teachers with feedback on lesson plans. These recommendations align with existing evidence showing that the most effective foundational learning programs encompass evidence-based support to teachers. Furthermore, given the potential role of teaching practices that foster student engagement in improving learning outcomes, including through child-centered and playful activities, it is important to explore how support for fostering such practices can be embedded in the design of training programs, and in teaching and learning materials such as lesson plans.

References

Akyeampong, Kwame, Tahir Andrabi, Abhijit Banerjee, Rukmini Banerji, Susan Dynarski, Rachel Glennerster, Sally Grantham-McGregor, Karthik Muralidharan, Benjamin Piper, Sara Ruto, Jaime Saavedra, Sylvia Schmelkes, and Hirokazu Yoshikawa. 2023. *Cost-Effective Approaches to Improve Global Learning - What does recent evidence tell us are "Smart Buys" for improving learning in low- and middle-income countries?* London, Washington D.C., New York. FCDO, the World Bank, UNICEF, and USAID. https://thedocs.worldbank.org/en/doc/231d98251cf326922518be0cbe306fdc-0200022023/related/GEEAP-Report-Smart-Buys-2023-final.pdf

ASER. 2023. Annual Status of Education Report (Rural) 2022: Provisional. https://asercentre.org/wp-content/uploads/2022/12/aserreport2022-1.pdf

Bazaldua, Diego Luna, Emma Jane Carter, and Laura Gregory. Forthcoming. Content Validity Study: Mapping the Lego Foundation's PLAY Classroom Observation Tool to the World Bank's Teach Primary.

Beisiegel, Mary, Rebecca Mitchell, & Heather C. Hill. 2018. The Design of Video-Based Professional Development: An Exploratory Experiment Intended to Identify Effective Features. *Journal of Teacher Education,* 69(1), 69-89. https://doi.org/10.1177/0022487117705096

Bietenbeck, Jan, Marc Piopiunik, and Simon Wiederhold. 2017. "Africa's Skill Tragedy: Does Teachers' Lack of Knowledge Lead to Low Student Performance?" *The Journal of Human Resources* 54: 857-899.

Bold, Tessa, Deon Filmer, Gayle Martin, Ezequiel Molina, Brian Stacy, Christophe Rockmore, Jakob Svensson, and Waly Wane. 2017. "Enrollment Without Learning: Teacher Effort, Knowledge, and Skill in Primary Schools in Africa." *Journal of Economic Perspectives* 31 (4): 185–204. https://doi.org/10.1257/jep.31.4.185

Bold, Tessa, Deon Filmer, Gayle Martin, Ezequiel Molina, Christophe Rockmore, Brian Stacy, Jakob Svensson, and Waly Wane. 2017. "What Do Teachers Know and Do? Does It Matter?: Evidence from Primary Schools in Africa." Policy Research Working Paper; No. 7956. © World Bank, Washington, DC. http://hdl.handle.net/10986/25964 License: CC BY 3.0 IGO.

Buhl-Wiggers, Julie, Jason T. Kerwin, Jeffrey A. Smith and Rebecca Thornton. 2017. The Impact of Teacher Effectiveness on Student Learning in Africa. RISE Programme, Oxford. https://riseprogramme.org/sites/default/files/inline-files/Buhl-Wiggers%20The%20Impact%20of%20Teacher%20Effectiveness%202017-04-30.pdf

Chi, Guanghua, Hang Fan, Sourav Chatterjee, and Joshua E. Blumenstock. 2021. "Microestimates of Wealth for All Low- and Middle-Income Countries." *Proceedings of the National Academy of Sciences* 119 (3). https://www.pnas.org/doi/full/10.1073/pnas.2113658119

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers. 2006. "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspectives* 20(1): 91–116.

Confemen. 2020. *Qualité des Systèmes Éducatifsen Afrique Subsaharienne Francophone.* https://www.confemen.org/wp-content/uploads/2022/07/Resume_Pasec2019_Web.pdf

Crawford, Michael, Maria Rebeca Barron Rodriguez, Elaine Yi Zhong Ding, Marcela Gutierrez Bernal, Jaime Saavedra Chanduvi, and Omas S. Arias Diaz. 2021. Attaining the Learning Target: A Policy Package to Promote Literacy for All Children (English). © World Bank, Washington, DC.

https://documents.worldbank.org/en/publication/documents-reports/documentdetail/565171627618840464/attaining-the-learning-target-a-policy-package-to-promote-literacy-for-all-children

Dobbie, Will, and Roland G. Fryer Jr. 2013. "Getting beneath the Veil of Effective Schools: Evidence from New York City." *American Economic Journal: Applied Economics,* 5 (4): 28-60.

Duflo, Esther, Rema Hanna, Stephen P. Ryan. 2012. Incentives work: getting teachers to come to school. *The American Economic Review*, 102(4), 1241–1278. https://doi.org/10.1257/aer.102.4.1241

Early Grade Reading Barometer. 2024. Published online at EarlyGradeReadingBarometer.com. Retrieved from: 'earlygradereadingbarometer.com'[Online Resource]

Evans, David K. and Tara Béteille. 2021. "Successful Teachers, Successful Students: Recruiting and Supporting Society's Most Crucial Profession." © World Bank, Washington, DC. https://thedocs.worldbank.org/en/doc/7b72550a6aec51f4e6fbbd31199d7b3b-0200022021/original/SuccessfulTeachers-Oct-5.pdf

Filmer, Deon, Ezequiel Molina, and Waly Wane. 2020. "Identifying Effective Teachers Lessons from Four Classroom Observation Tools." Policy Research Working Paper, No. 9365. © World Bank, Washington, DC. https://elibrary.worldbank.org/doi/epdf/10.1596/1813-9450-9365.

Filmer, Deon P.; Nahata, Vatsal; Sabarwal, Shwetlena. 2021. "Preparation, Practice, and Beliefs : A Machine Learning Approach to Understanding Teacher Effectiveness". Policy Research working paper,no. WPS 9847 Washington, D.C. : World Bank Group. http://documents.worldbank.org/curated/en/737101636989604541/Preparation-Practice-and-Beliefs-A-Machine-Learning-Approach-to-Understanding-Teacher-Effectiveness

Filmer, Deon P., Vatsal Nahata, and Shwetlena Sabarwal. 2021. "Preparation, Practice, and Beliefs: A Machine Learning Approach to Understanding Teacher Effectiveness," Policy Research Working Paper Series 9847, The World Bank.

Global Education Policy Dashboard. 2024. Published online at EducationPolicyDashboard.com. Retrieved from: 'https://www.educationpolicydashboard.org/' [Online Resource]

Jukes, Matthew. 2023. "New Toolkit Measures How Teachers and Caregivers Support Children's Engagement in Learning." Published online at shared.rti.org. Retrieved from: 'https://shared.rti.org/content/new-toolkit-measures-how-teachers-and-caregivers-support-children%E2%80%99s-engagement-learning' [Online Resource]

Jukes, M. C. H., Yoshikawa, H., Betts, K., Dubeck, M., Edwards, L., Nduku, T., Staskowicz, E., Stern, J., Gjicali, K., Kim, S., Mahbub, T., Moran, C., Patankar, K. U., Rosenbach, S., Saleh, H. M., Strouf, K., & Zhao, V. Y. 2022. *Playful Learning Across the Years (PLAY) Measurement Initiative: Full Report*.

Lavy, Victor, 2015. "Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries." *The Economic Journal*, vol 125(588), F397-F424.

Le Nestour, Alexis, Laura Moscoviz, and Justin Sandefur. 2022. "The Long-Run Decline of Education Quality in the Developing World". Center for Global Development, Washington, DC. https://www.cgdev.org/publication/long-run-decline-education-quality-developing-world

Metzler, Johannes, and Ludger Woessmann. 2012. "The impact of teacher subject knowledge on student achievement: Evidence from within-teacher within-student variation." *Journal of Development Economics* 99, 486–496.

Molina, Ezequiel, Syeda Farwa Fatima, Andrew Dean Ho, Carolina Melo, Tracy Marie Wilichowski, and Adelle Pushparatnam. 2020. "Measuring the Quality of Teaching Practices in Primary Schools: Assessing the Validity of the Teach Observation Tool in Punjab, Pakistan." *Teaching and Teacher Education* 96: 103171.

Muijs, Daniel, Leonidas Kyriakides, Greetje van der Werf, Bert Creemers, Helen Timperley, and Lorna Earl. 2014. "State of the Art—Teacher Effectiveness and Professional Learning." *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice* 25(2): 231–56.

Ritchie, Hannah, Veronika Samborska, Natasha Ahuja, Esteban Ortiz-Ospina and Max Roser. 2023. "Global Education" Published online at OurWorldInData.org. Retrieved from: https://ourworldindata.org/global-education [Online Resource]

Snilstveit, B, Jennifer Stevenson, Daniel Phillips, Martina Vojtkova, Emma Gallagher, Tanja Schmidt, Hannah Jobse, Maisie Geelen, Maria Grazia Pastorello, M, and John Eyers. 2015. *Interventions for improving learning outcomes and access to education in low- and middle- income countries: a systematic review, 3ie Systematic Review 24.* London: International Initiative for Impact Evaluation (3ie). https://www.3ieimpact.org/sites/default/files/2019-01/SR24-education-review_2.pdf

Stern, Jonathan, Matthew Jukes, Benjamin Piper, Joseph DeStefano, Jessica Mejia, Peggy Dubeck, Bidemi Carrol, Rachel Jordan, Cosmus Gatuyu, Tabitha Nduku, Maitri Punjabi, Christine Harris Van Keuren, and Fatima Tufail. 2021. Learning at Scale: Interim Report. RTI International. https://ierc-publicfiles.s3.amazonaws.com/public/resources/Learning%20at%20Scale%20Interim%20Report%20-%20Final%20Draft.pdf

World Bank. 2022. *The State of Global Learning Poverty: 2022 Update.* © World Bank, Washington, DC. https://thedocs.worldbank.org/en/doc/e52f55322528903b27f1b7e61238e416-0200022022/original/Learning-poverty-report-2022-06-21-final-V7-0-conferenceEdition.pdf

Annex A – Descriptive Tables by Education System

Table A1: Teacher Characteristics

		Low-income					Lower-middle income						Upper-middle income	
	All	Ethiopia	Niger	Rwanda	Sierra Leone	Chad	Jordan	Pakistan ICT	Pakistan KP	Pakistan Balochistan	Pakistan Sindh	Pakistan Punjab	Peru	Gabon
Number of teachers	2686	287	262	194	240	240	246	100	197	195	185	200	186	154
Age	38.9	33.4	35.7	36.5	37.5	36.8	36.6	39.3	33.7	41.8	42.6	39.7	47.7	43.6
	(.5)	(1.4)	(.9)	(1.2)	(1.5)	(1.2)	(1.4)	(2.7)	(.6)	(.9)	(1)	(1)	(1.9)	(4.4)
	[10.4]	[8.4]	[7.9]	[9.2]	[9.6]	[9.7]	[7.9]	[10.5]	[6.9]	[8.3]	[11.3]	[10.6]	[9.9]	[12.4]
Teaching experience (in years)	13.6	14.9	10.6	11.4	12	10.9	9.3	13.6	7.6	18.6	15.2	14.5	21.5	15.4
	(.6)	(1.7)	(.6)	(1.3)	(1.4)	(.8)	(1)	(2.2)	(.6)	(1.1)	(1.2)	(1)	(1.9)	(5.1)
	[10.3]	[9.9]	[5.5]	[9.5]	[8]	[7.4]	[6.1]	[9.9]	[7.3]	[10.4]	[13.5]	[10.6]	[9.8]	[13.4]
% male	53	65.8	57.2	48.2	84	12.4	86.6	64	38.3	48.1	28.1	53.2	77.4	31.6
	(2.1)	(7.1)	(5.4)	(6.4)	(5.1)	(3.7)	(5.7)	(10.8)	(4.3)	(4.8)	(4.3)	(4.3)	(7.8)	(14.3)
	[49.9]	[47.6]	[49.6]	[50.3]	[36.9]	[33.1]	[34.2]	[48.3]	[48.7]	[50.1]	[45.1]	[50]	[42]	[46.7]
No education	.1	0	0	0	0	0	0	0	0	0	0	.7	0	0
	(.1)	(.)	(.)	(.)	(.)	(0)	(.)	(.)	(.)	(.)	(.)	(.7)	(.)	(.)
	[2.6]	[0]	[0]	[0]	[0]	[1.6]	[0]	[0]	[0]	[0]	[0]	[8.6]	[0]	[0]
Less than primary	0	0	.1	0	0	0	0	0	0	0	0	0	0	0
	(0)	(.)	(.1)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
	[1]	[0]	[3.3]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]
Primary completed	.3	.1	.1	0	.1	.8	0	0	0	0	0	0	0	2.4
	(.2)	(.1)	(.1)	(.)	(.1)	(.6)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(2.4)
	[5.3]	[2.5]	[3.5]	[0]	[3.1]	[8.8]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[15.3]
Middle passed	4	0	34.2	.7	9.3	4	0	0	0	0	0	0	0	0
	(.7)	(0)	(6)	(.7)	(4.4)	(1.6)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(0)
	[19.6]	[2.2]	[47.6]	[8.7]	[29.3]	[19.7]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[.9]
High school passed	22.2	1.8	13.8	90.7	64.8	22.2	0	0	3.1	21	9.7	15.1	.1	26.9
	(1.8)	(.8)	(3.4)	(3.6)	(7.3)	(5.3)	(.)	(.)	(1.5)	(4.3)	(2.9)	(3.7)	(.1)	(14.3)
	[41.6]	[13.3]	[34.6]	[29.2]	[48]	[41.7]	[0]	[0]	[17.3]	[40.9]	[29.7]	[35.9]	[2.6]	[44.6]
Bachelor's degree	32.7	18.2	6.4	5.8	4.2	42.9	79	20.8	20.7	50.8	58	17.4	56.4	61.8
	(2.3)	(6.8)	(2.5)	(3)	(2.3)	(6.5)	(7.6)	(13.7)	(4.1)	(5.4)	(4.6)	(3.5)	(10.2)	(15.1)
	[46.9]	[38.7]	[24.6]	[23.6]	[20.3]	[49.7]	[40.9]	[41]	[40.6]	[50.2]	[49.5]	[38.1]	[49.8]	[48.9]
Master's degree	22.3	0	0	0	0	2.1	6.1	78.4	74.4	27.2	32.2	63.4	21.6	2.5
	(1.5)	(.)	(.)	(.)	(.)	(1.5)	(3.7)	(13.7)	(4.4)	(4.5)	(4.3)	(4.5)	(9.6)	(2.2)
	[41.6]	[0]	[0]	[0]	[0]	[14.5]	[24.2]	[41.5]	[43.8]	[44.6]	[46.9]	[48.3]	[41.3]	[15.6]
Doctoral degree	.7	0	0	0	0	0	0	.8	.9	0	.1	3.2	3.9	0
	(.4)	(0)	(.)	(.)	(.)	(.)	(.)	(.8)	(.9)	(.)	(.1)	(1.4)	(3.9)	(.)
	[8.4]	[1.3]	[0]	[0]	[0]	[0]	[0]	[8.8]	[9.4]	[0]	[3.4]	[17.8]	[19.5]	[0]
Teachers Training Institute (TTI) Graduate	.6	7.1	0	0	0	0	0	0	0	0	0	0	0	0
	(.4)	(4.4)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)

	[7.7]	[25.8]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]
Teachers Training College (TTC) Graduate	5.2	62.5	0	0	0	0	0	0	0	0	0	0	0	0
	(.9)	(7.7)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
	[22.3]	[48.6]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]	[0]
Other	11.9	10.3	45.3	2.7	21.5	27.9	14.8	0	1	1.1	0	0	18.1	6.5
	(1.3)	(4.2)	(5.7)	(2)	(6.2)	(6)	(6.8)	(.)	(.7)	(.7)	(.)	(.)	(7.6)	(5.3)
	[32.4]	[30.5]	[49.9]	[16.3]	[41.3]	[45]	[35.7]	[0]	[10.1]	[10.3]	[0]	[0]	[38.6]	[24.9]
Language	19.9	42.5	1.9	67.1	.1	1	50.4	50	7.4	19.1	13.7	11.3	0	0
	(1.6)	(7.2)	(1.6)	(5.9)	(.1)	(.7)	(8.2)	(9.6)	(2.5)	(3.2)	(2.7)	(2.4)	(.)	(.)
	[39.9]	[49.6]	[13.7]	[47.3]	[2.4]	[10.2]	[50.2]	[50.3]	[26.2]	[39.4]	[34.5]	[31.7]	[0]	[0]
Mathematics	9.3	41.4	0	23	4.1	3	19.9	17.1	2.6	8.9	1.8	1	0	.1
	(1)	(6.9)	(.)	(5.4)	(2.9)	(3)	(5.3)	(5.2)	(1.2)	(2.9)	(1)	(.4)	(.)	(.1)
	[29]	[49.4]	[0]	[42.4]	[19.9]	[17.2]	[40.1]	[37.9]	[16.1]	[28.5]	[13.5]	[9.8]	[0]	[3.4]
Both/All	68.8	14.4	98.1	9.9	95.9	95.8	13.3	32.5	88.5	65.1	84.1	87	99.2	99.9
	(1.8)	(6.3)	(1.6)	(3.2)	(2.9)	(3)	(6.5)	(9)	(2.9)	(4.5)	(2.9)	(2.5)	(.8)	(.1)
	[46.3]	[35.2]	[13.7]	[30.1]	[20]	[20.2]	[34.1]	[47.1]	[32]	[47.8]	[36.6]	[33.7]	[8.8]	[3.4]
Other	2	1	0	0	0	.1	16.4	.4	1.5	6.9	.3	.7	.8	0
	(.5)	(.7)	(.)	(.)	(.)	(.1)	(5.9)	(.4)	(.8)	(3.1)	(.3)	(.4)	(.8)	(.)
	[14]	[9.9]	[0]	[0]	[0]	[3.7]	[37.2]	[6.4]	[12.3]	[25.4]	[5.8]	[8.4]	[8.8]	[0]

Sample is restricted to teachers selected for the classroom observation module. Table shows average with standard error in parentheses and standard deviations in brackets. Values in the overall column are the unweighted average across countries. Values in country columns are weighted. Two categories, Teachers Training Institute (TTI) Graduate and Teachers Training College (TTC) Graduate, were only included in the survey instrument in Ethiopia.

Table A2: Student Characteristics

		Low-income					Lower-middle income						Upper-middle income	
	All	Ethiopia	Niger	Rwanda	Sierra Leone	Chad	Jordan	Pakistan ICT	Pakistan KP	Pakistan Balochistan	Pakistan Sindh	Pakistan Punjab	Peru	Gabon
Number of students	50392	5783	5074	5077	4073	4988	4932	2420	3768	2479	2350	3972	2517	2959
Age	10.6	11.6	11	11.1	10.9	11.6	9.5	10.4	10.4	10.9	10.7	10.3	9.6	10
	(0)	(.1)	(.2)	(0)	(.)	(.1)	(.1)	(.2)	(.1)	(.2)	(.1)	(.1)	(.)	(.2)
	[1.6]	[1.8]	[1.4]	[1.7]	[1.7]	[1.8]	[.7]	[1.4]	[1.3]	[1.6]	[1.5]	[1.3]	[.7]	[1.6]
Household size	6.7	5.7	7.6	5.7	.	7.5	6.2	6.8	7.3	7.5	7.4	6.8	.	5.8
	(0)	(.1)	(.1)	(.1)	(.1)	(.1)	(0)	(.1)	(.1)	(.1)	(.1)	(.1)	(0)	(.2)
	[2.3]	[2.1]	[2.5]	[1.8]	[.]	[2.6]	[1.7]	[2.1]	[2.1]	[2.6]	[2.1]	[2]	[.]	[2.5]
Parents/caregivers went to university	46.6	31.7	40.4	37.8	.	54.9	66.7	47.6	34.6	35.2	29.9	43.3	.	83.4
	(1.4)	(3.7)	(6.2)	(1.8)	(.)	(3.1)	(2.1)	(2.8)	(2.7)	(3.2)	(4.2)	(2.3)	(.)	(3.3)
	[49.9]	[46.5]	[49.1]	[48.5]	[.]	[49.8]	[47.1]	[50]	[47.6]	[47.8]	[45.8]	[49.6]	[.]	[37.2]
Student did not eat yesterday	.5	3.3	.4	.1	.	.2	.4	.1	0	0	.1	0	.	.1
	(.1)	(1.2)	(.1)	(0)	(.)	(.1)	(.2)	(.1)	(0)	(.)	(.1)	(.)	(.)	(.1)
	[6.8]	[18]	[6.3]	[2.5]	[.]	[4.4]	[6.1]	[3.3]	[2.1]	[0]	[3.6]	[0]	[.]	[3.2]
Student ate one to two times a day	41.9	19.2	36.7	63.7	.	43.3	39	35.7	40	59.1	52.3	31.3	.	42.7
	(.9)	(1.3)	(3.8)	(1.4)	(.)	(3.1)	(1.1)	(4.3)	(2.8)	(3.6)	(3.2)	(1.6)	(.)	(3.8)
	[49.3]	[39.4]	[48.2]	[48.1]	[.]	[49.6]	[48.8]	[47.9]	[49]	[49.2]	[50]	[46.4]	[.]	[49.5]
Student ate at least 3 times a day	57.7	77.5	62.9	36.3	.	56.5	60.7	64.2	60	40.9	47.6	68.7	.	57.2
	(.9)	(1.8)	(3.8)	(1.4)	(.)	(3.1)	(1.1)	(4.3)	(2.8)	(3.6)	(3.2)	(1.6)	(.)	(3.8)
	[49.4]	[41.8]	[48.3]	[48.1]	[.]	[49.6]	[48.9]	[48]	[49]	[49.2]	[50]	[46.4]	[.]	[49.5]
Student has 0 pairs of shoes	.5	4	.3	0	.	.3	.4	0	.2	0	0	0	.	0
	(.2)	(1.5)	(.1)	(0)	(.)	(.1)	(.2)	(0)	(.1)	(.)	(.)	(.)	(.)	(0)
	[7.2]	[19.6]	[5.6]	[1.9]	[.]	[5]	[6.5]	[1.2]	[4]	[0]	[0]	[0]	[.]	[2]
Student has 1 pair of shoes	21.3	15	21.2	18.7	.	17.8	9.9	14.4	29.3	41.5	44.4	21.1	.	4.4
	(.9)	(1.6)	(4.3)	(.9)	(.)	(1.8)	(1.2)	(3)	(2.4)	(3)	(3)	(1.4)	(.)	(.9)
	[40.9]	[35.7]	[40.9]	[39]	[.]	[38.2]	[29.9]	[35.1]	[45.5]	[49.3]	[49.7]	[40.8]	[.]	[20.4]
Student has at least 2 pairs of shoes	78.2	81	78.4	81.3	.	82	89.6	85.6	70.5	58.5	55.6	78.9	.	95.6
	(.9)	(2)	(4.3)	(.9)	(.)	(1.8)	(1.2)	(3)	(2.4)	(3)	(3)	(1.4)	(.)	(.9)
	[41.3]	[39.3]	[41.1]	[39]	[.]	[38.5]	[30.5]	[35.1]	[45.6]	[49.3]	[49.7]	[40.8]	[.]	[20.5]
Bathroom at home: Toilet flush	24	5.8	6.3	9.1	.	6.9	71	33.9	21	13.7	11.5	23	.	62.7
	(1.2)	(.9)	(1)	(.8)	(.)	(1.4)	(1.8)	(2.3)	(1.6)	(1.8)	(1.4)	(1.4)	(.)	(6)
	[42.7]	[23.4]	[24.3]	[28.8]	[.]	[25.3]	[45.4]	[47.4]	[40.7]	[34.4]	[31.9]	[42.1]	[.]	[48.4]
Bathroom at home: Pit	28.2	43.5	30.5	75.3	.	17.8	16.3	21.1	15.8	34.6	24.1	25.1	.	5.7
	(1)	(2.4)	(4.6)	(1.3)	(.)	(2.2)	(1.5)	(2)	(1.6)	(3.3)	(3.4)	(1.8)	(.)	(1.8)
	[45]	[49.6]	[46.1]	[43.1]	[.]	[38.2]	[36.9]	[40.8]	[36.5]	[47.6]	[42.8]	[43.4]	[.]	[23.2]
Bathroom at home: Bucket	5.1	5.9	2.4	1	.	3.4	.1	6.8	6.7	8.2	10.6	8.7	.	2.2
	(.4)	(1.7)	(.5)	(.2)	(.)	(1.3)	(.1)	(1.1)	(.9)	(1.2)	(1.8)	(.9)	(.)	(.8)
	[22]	[23.5]	[15.4]	[10.1]	[.]	[18.1]	[2.8]	[25.3]	[25]	[27.5]	[30.8]	[28.1]	[.]	[14.5]

Bathroom at home: Communal bathroom	21.1	20	14.8	12.4	.	24.7	9.2	23.8	26.6	27.5	25	36.5	.	11.9
	(.7)	(2.1)	(1.6)	(1.1)	(.)	(3.2)	(.9)	(1.2)	(2.3)	(2)	(2.6)	(1.8)	(.)	(2.3)
	[40.8]	[40]	[35.5]	[33]	[.]	[43.1]	[28.9]	[42.6]	[44.2]	[44.6]	[43.3]	[48.2]	[.]	[32.4]
No bathroom at home (outside)	7.3	7.2	26.9	.4	.	16.9	.2	3.4	4.9	7.3	5.6	2.8	.	4.3
	(.6)	(1.1)	(3.5)	(.1)	(.)	(2.2)	(.1)	(1.7)	(1.4)	(1.7)	(1)	(.5)	(.)	(2.1)
	[26]	[25.9]	[44.4]	[6.6]	[.]	[37.5]	[4.4]	[18.2]	[21.6]	[26.1]	[23]	[16.6]	[.]	[20.3]
Bathroom at home: No response	14.3	17.6	19	1.6	.	30.4	3.3	10.9	25	8.7	23.3	3.8	.	13.2
	(.9)	(2.4)	(2.4)	(.3)	(.)	(3.7)	(.7)	(1.7)	(3)	(1.6)	(5.2)	(.6)	(.)	(3.6)
	[35]	[38.1]	[39.2]	[12.7]	[.]	[46]	[17.8]	[31.1]	[43.3]	[28.2]	[42.3]	[19.2]	[.]	[33.9]

Sample is restricted to students assessed for student learning on literacy and mathematics. Table shows average with standard error in parentheses and standard deviations in brackets. Values in the overall column are the unweighted average across countries. Values in the overall column show the average of the values across countries. Information about household size, if parents went to university, number of times student ate per day, number of pairs of shoes they own, and the type of bathroom they have at home were not asked in Sierra Leone and Peru.

Table A3: Student Knowledge

		Low-income					Lower-middle income						Upper-middle income	
	All	Ethiopia	Niger	Rwanda	Sierra Leone	Chad	Jordan	Pakistan ICT	Pakistan KP	Pakistan Balochistan	Pakistan Sindh	Pakistan Punjab	Peru	Gabon
Number of students	50392	5783	5074	5077	4073	4988	4932	2420	3768	2479	2350	3972	2517	2959
Student knowledge	52.5	55.1	40.2	40.1	50.1	36	62.8	54.3	44.3	49.4	47.4	55.9	80.7	66.5
	(.8)	(1.6)	(1.5)	(.9)	(1.8)	(.9)	(1.2)	(1.9)	(1.7)	(1.9)	(2.1)	(1)	(1.3)	(3)
	[22.2]	[21.7]	[18.9]	[17.2]	[20.6]	[16.5]	[17.2]	[16.5]	[17.9]	[23.9]	[20.5]	[18.7]	[14.4]	[18.8]
Literacy student knowledge	64.7	67.5	42.9	51.8	63.8	40.3	85	69.1	55.1	60.5	63.5	69	92.4	80.8
	(.8)	(1.7)	(1.4)	(1.6)	(2)	(1.3)	(1.1)	(2.6)	(1.8)	(2.1)	(2.6)	(1)	(1.1)	(2.2)
	[26.8]	[25.6]	[21.9]	[24.5]	[25]	[20.3]	[19.8]	[21.6]	[21.9]	[27.1]	[24.5]	[22.2]	[13.2]	[20.4]
Mathematics student knowledge	40.3	42.7	37.5	28.5	36.5	31.8	40.6	39.4	33.5	38.3	31.3	42.9	69.1	52.1
	(.8)	(1.8)	(1.8)	(.7)	(2.1)	(1.2)	(1.6)	(1.6)	(1.8)	(2)	(2.1)	(1.2)	(1.9)	(4.3)
	[24.1]	[23.8]	[21.1]	[18]	[22]	[22.1]	[21]	[17.8]	[20]	[25.9]	[23.4]	[21.1]	[20.8]	[25.5]
Pupil can link pictures	81.2	87.5	41.1	73.1	74.3	42.9	96.3	94.4	82.9	87.3	89.4	94.3	99.4	93.3
	(.7)	(1.4)	(2)	(2)	(2.2)	(1.4)	(.7)	(1.1)	(1.9)	(2)	(2)	(.6)	(.2)	(1.1)
	[32.4]	[27.8]	[38.5]	[36.5]	[36.8]	[24.2]	[15]	[17.7]	[30.4]	[28.4]	[26.8]	[19.3]	[6.2]	[16.9]
Pupil can understand a story (1)	51.9	60.6	25.4	27.7	50.3	21.6	77.9	50.6	31.7	47.5	50.1	53.9	95	83
	(1.4)	(2.6)	(4.3)	(2.3)	(3.4)	(2.6)	(1.8)	(5.4)	(3.6)	(3.9)	(4.4)	(2.5)	(1.4)	(2)
	[50]	[48.9]	[43.5]	[44.8]	[50]	[41.2]	[41.5]	[50]	[46.5]	[49.9]	[50]	[49.9]	[21.8]	[37.6]
Pupil can understand a story (2)	52.5	62.2	27.9	33.1	51.8	24.4	85.1	52.8	24	43.3	49	51.2	90.2	87.9
	(1.4)	(2.6)	(4.4)	(2.1)	(3.3)	(2.5)	(1.6)	(6.2)	(3)	(3.7)	(4.4)	(2.5)	(1.8)	(1.8)
	[49.9]	[48.5]	[44.9]	[47.1]	[50]	[43]	[35.7]	[49.9]	[42.7]	[49.6]	[50]	[50]	[29.7]	[32.6]
Pupil can select a correct word	33.4	36.7	9	12.7	34.4	11.3	64.7	32.8	18.4	27.1	19.4	30.3	79.5	58.3
	(1.2)	(3.1)	(1.5)	(1)	(3.5)	(1.9)	(2.2)	(2.1)	(2.2)	(2.7)	(2.4)	(1.6)	(1.9)	(2.3)
	[47.2]	[48.2]	[28.7]	[33.4]	[47.5]	[31.7]	[47.8]	[47]	[38.8]	[44.5]	[39.5]	[45.9]	[40.4]	[49.3]
Pupil can recognize a letter	83.6	77.8	74.4	68.3	91.6	72.8	92.1	95.6	89	77.4	79.7	91.5	94.1	82.3
	(.8)	(3.3)	(2.9)	(3)	(1.7)	(2.2)	(1.5)	(.6)	(1.4)	(2.5)	(4.3)	(.8)	(1.1)	(4.6)
	[30.9]	[35.4]	[35.2]	[38.8]	[20.2]	[37.1]	[22.2]	[15.9]	[25.6]	[35.4]	[33.6]	[22.5]	[17]	[32]
Pupil can recognize a word	85.6	79.9	79.3	95.6	80.4	68.6	93.7	88.5	84.4	79.8	93.7	92.5	96.1	79.9
	(.8)	(2.8)	(2.9)	(1.1)	(2.6)	(2.5)	(1.1)	(2.1)	(1.5)	(2.1)	(1.1)	(.9)	(1.7)	(5.9)
	[29.6]	[35.2]	[29.9]	[20.1]	[31.4]	[40.2]	[21.1]	[25.4]	[29.4]	[30.9]	[20]	[20.5]	[17.3]	[37.1]
Pupil can recognize numbers	88.9	84.3	83.6	85.2	93.1	80.4	90.6	96	92.9	83.6	91.8	91.3	98.3	84.6
	(.6)	(2.6)	(3.2)	(1.7)	(1.2)	(2.4)	(1.3)	(.8)	(.8)	(1.8)	(1.2)	(.9)	(.3)	(4.7)
	[27.1]	[32.9]	[30.2]	[29.4]	[19.1]	[36.1]	[24.2]	[15.9]	[21.4]	[31.4]	[24]	[23]	[9.8]	[34.3]
Pupil can order numbers	42.1	52.9	59.7	24.2	33.5	15.4	62.8	51.1	28.7	34.1	24.6	33.2	81.3	45.4
	(1.1)	(2.3)	(3.1)	(1.3)	(3.3)	(2.1)	(2.5)	(4.2)	(2.5)	(3)	(2.9)	(2.2)	(1.6)	(3.2)
	[49.4]	[49.9]	[49.1]	[42.8]	[47.2]	[36.1]	[48.3]	[50]	[45.2]	[47.4]	[43.1]	[47.1]	[39]	[49.8]
Pupil can add single digits	76	78.4	72.2	66.2	75.5	70.7	70.3	82.7	73.8	71.6	62.5	85.1	94.1	84.8
	(.9)	(1.9)	(2.7)	(1.4)	(2.1)	(2.6)	(2.7)	(2.2)	(2.2)	(2.7)	(4.6)	(1)	(1)	(3.5)

	[42.7]	[41.1]	[44.8]	[47.3]	[43]	[45.5]	[45.7]	[37.9]	[44]	[45.1]	[48.4]	[35.6]	[23.6]	[35.9]
Pupil can add double digits	61.3	67.3	59	46	52.5	52.1	65.1	65.1	55.9	49.8	47.2	72.3	90.1	74.5
	(1)	(1.9)	(3.3)	(1.3)	(3.3)	(2.2)	(2.4)	(2.6)	(2.3)	(3.5)	(3.9)	(1.6)	(1.5)	(4)
	[48.7]	[46.9]	[49.2]	[49.8]	[49.9]	[50]	[47.7]	[47.7]	[49.7]	[50]	[49.9]	[44.8]	[29.8]	[43.6]
Pupil can add triple digits	56.2	54.2	60.8	48.8	39.2	38.8	58	63.2	49.5	45.7	43.7	66.1	91	71.9
	(1.1)	(2.1)	(2.8)	(1.5)	(3.1)	(2.4)	(2.3)	(2.6)	(2.7)	(3.3)	(3.7)	(1.8)	(1.8)	(4.6)
	[49.6]	[49.8]	[48.8]	[50]	[48.8]	[48.7]	[49.4]	[48.2]	[50]	[49.8]	[49.6]	[47.3]	[28.6]	[45]
Pupil can subtract single digits	69.2	68.9	67.7	63.6	66	60.3	74.3	74.3	62.2	63.5	52.6	75	89.6	81.2
	(.9)	(2.1)	(2.8)	(1.4)	(2.7)	(2.9)	(2.5)	(2.4)	(2.4)	(3)	(4.2)	(1.5)	(1)	(4.1)
	[46.2]	[46.3]	[46.8]	[48.1]	[47.4]	[48.9]	[43.7]	[43.7]	[48.5]	[48.1]	[49.9]	[43.3]	[30.6]	[39.1]
Pupil can subtract double digits	39.9	41.8	40.1	24.3	30.7	25.9	40.2	46.5	30.2	36.3	29.8	50.2	64.1	58.8
	(1.1)	(2.1)	(3.8)	(1.3)	(3.7)	(2.1)	(2.7)	(3.9)	(2.1)	(3.5)	(2.8)	(2.3)	(2.4)	(7)
	[49]	[49.3]	[49]	[42.9]	[46.1]	[43.8]	[49]	[49.9]	[45.9]	[48.1]	[45.7]	[50]	[48]	[49.2]
Pupil can multiply single digits	38.6	39.9	37.7	25	42.9	47.7	30.3	30.7	25.6	29.4	21.9	39.7	74	57.5
	(1.2)	(2.1)	(3.6)	(1.5)	(3.3)	(2.7)	(2.1)	(3.2)	(2.9)	(3.3)	(2.3)	(1.9)	(2.3)	(6.1)
	[48.7]	[49]	[48.5]	[43.3]	[49.5]	[50]	[45.9]	[46.1]	[43.7]	[45.5]	[41.4]	[48.9]	[43.9]	[49.4]
Pupil can multiply double digits	13.8	14.5	15.1	4.3	5.6	8.8	3	7.3	4.9	11.3	8.7	10.8	45.4	40
	(1.1)	(2.4)	(1.8)	(.6)	(1.7)	(1)	(.5)	(.9)	(2.3)	(2.1)	(1.3)	(1.1)	(3.9)	(6.8)
	[34.5]	[35.3]	[35.8]	[20.2]	[23]	[28.4]	[17.1]	[26]	[21.6]	[31.6]	[28.3]	[31.1]	[49.8]	[49]
Pupil can multiply triple digits	10	10.2	7.5	2.6	4.6	5.7	1.5	9.2	1.9	8.6	6.1	8.4	39.2	24.7
	(.9)	(1.8)	(1.1)	(.5)	(1.6)	(.8)	(.4)	(.9)	(1)	(2)	(1.2)	(.9)	(3.4)	(6.2)
	[30]	[30.3]	[26.3]	[15.8]	[20.9]	[23.2]	[12.3]	[28.9]	[13.7]	[28]	[23.9]	[27.8]	[48.8]	[43.1]
Pupil can divide single digits	36.8	39.7	26.9	19.8	33.6	34	35.9	34	29.1	38.2	21.5	42.5	73	50.3
	(1.3)	(2.7)	(4.4)	(1)	(4.2)	(2.7)	(2.5)	(3.4)	(3.4)	(3)	(2.9)	(2.4)	(3.4)	(9.1)
	[48.2]	[48.9]	[44.3]	[39.9]	[47.2]	[47.4]	[48]	[47.4]	[45.4]	[48.6]	[41.1]	[49.4]	[44.4]	[50]
Pupil can divide double digits	22.7	27.1	14.6	10.3	25.2	13.1	14.9	14.5	18.4	20.9	12.2	28.6	54.7	40.6
	(1.2)	(2.5)	(4.4)	(.8)	(4.3)	(1.4)	(2)	(1.8)	(3.4)	(2.5)	(1.9)	(2.3)	(3.7)	(6.9)
	[41.9]	[44.4]	[35.3]	[30.4]	[43.4]	[33.7]	[35.6]	[35.3]	[38.8]	[40.7]	[32.7]	[45.2]	[49.8]	[49.1]
Pupil understands division	12.9	17.7	7	2.4	6	8.6	18.5	3.9	7.4	26.7	11.1	8.4	34.8	15.4
	(.8)	(2.6)	(1.8)	(.4)	(1.3)	(1.1)	(2.3)	(.8)	(1.2)	(3.6)	(1.6)	(.9)	(3.5)	(2.3)
	[33.6]	[38.1]	[25.6]	[15.4]	[23.8]	[28.1]	[38.9]	[19.3]	[26.2]	[44.3]	[31.4]	[27.8]	[47.7]	[36.1]
Pupil can solve math story	14.5	15.8	4.3	2.3	14.3	7	17.1	2.4	6.5	21.9	14.7	14.9	45.5	22.3
	(.9)	(2.1)	(1.3)	(.5)	(3.8)	(.9)	(2.3)	(.5)	(2)	(2.7)	(2.2)	(2)	(3.6)	(3.4)
	[35.2]	[36.5]	[20.4]	[15.1]	[35]	[25.5]	[37.7]	[15.2]	[24.6]	[41.4]	[35.5]	[35.6]	[49.8]	[41.6]
Pupil can complete a sequence	21.6	28.4	6.4	2.6	24.2	8.1	26.5	10.8	16	29.6	20.9	16.5	61.3	30.1
	(1.2)	(2.8)	(1.7)	(.4)	(4.6)	(1)	(2.6)	(1)	(3.6)	(3.2)	(2.5)	(1.5)	(3.3)	(3.7)
	[41.2]	[45.1]	[24.4]	[16]	[42.8]	[27.3]	[44.1]	[31]	[36.6]	[45.7]	[40.7]	[37.1]	[48.7]	[45.9]

Sample is restricted to students assessed for student learning on literacy and mathematics. Table shows average with standard error in parentheses. Values in the overall column are the unweighted average across countries. Student knowledge score is the average of literacy and mathematics knowledge scores. Literacy student knowledge score measures the average score each student received across the 24 questions in the assessment. Mathematics student knowledge score measures the average score each student received across 15 mathematics questions. Values in the overall column are unweighted. Values in country columns are weighted.

Table A4: Teacher Pedagogical Skills

		Low-income					Lower-middle income						Upper-middle income	
	All	Ethiopia	Niger	Rwanda	Sierra Leone	Chad	Jordan	Pakistan ICT	Pakistan KP	Pakistan Balochistan	Pakistan Sindh	Pakistan Punjab	Peru	Gabon
Number of teachers	2686	287	262	194	240	240	246	100	197	195	185	200	186	154
TEACH score	2.7	2.4	2.5	2.7	2.4	2.5	3.2	3.2	2.6	2.6	2.6	2.6	3.1	3.3
	(0)	(.1)	(0)	(0)	(0)	(.1)	(.1)	(.1)	(0)	(.1)	(.1)	(0)	(.1)	(.2)
	[.7]	[.5]	[.3]	[.3]	[.3]	[.6]	[.9]	[.5]	[.5]	[.7]	[.8]	[.5]	[.6]	[.6]
Supportive learning environment	3.3	3	3.3	3.4	2.9	3.3	3.3	3.6	3.3	3.1	2.9	3.2	4	4
	(0)	(.1)	(.1)	(.1)	(.1)	(.1)	(.2)	(.1)	(.1)	(0)	(.1)	(0)	(.1)	(.2)
	[.8]	[.7]	[.6]	[.6]	[.7]	[.8]	[1.1]	[.7]	[.6]	[.6]	[.7]	[.5]	[.5]	[.6]
Positive behavioral expectations	3	2.4	3	3.4	2.6	2.5	3.1	3.5	2.8	2.8	2.7	2.7	3.5	3.6
	(0)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.2)
	[.9]	[.8]	[.6]	[.7]	[.7]	[1]	[1]	[.7]	[.8]	[.8]	[1]	[.8]	[.8]	[.9]
Lesson facilitation	3.1	2.9	2.8	3.3	3	3.1	3.5	3.6	2.8	2.5	2.7	2.8	3.3	3.4
	(0)	(.1)	(.1)	(.1)	(.1)	(.1)	(.2)	(.2)	(.1)	(.1)	(.1)	(.1)	(.2)	(.3)
	[.9]	[.6]	[.7]	[.7]	[.5]	[.8]	[1]	[.9]	[1]	[.9]	[1]	[.9]	[.9]	[.8]
Checks for understanding	2.9	2.9	2.6	2.9	2.5	2.5	3.4	3.3	2.8	2.5	2.5	2.7	3.2	3.9
	(0)	(.1)	(.1)	(.1)	(.1)	(.1)	(.2)	(.1)	(.1)	(.1)	(.1)	(.1)	(.2)	(.3)
	[1]	[.9]	[.8]	[.9]	[.5]	[1]	[1.1]	[.7]	[.8]	[.8]	[.9]	[.7]	[1]	[.9]
Feedback	2.5	1.9	2.1	1.6	2	2.8	3.3	3	2.1	2.6	2.7	2.6	2.4	4
	(0)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.2)	(.2)
	[1.1]	[1]	[.9]	[.5]	[.6]	[1.2]	[1]	[1]	[.8]	[.9]	[1.2]	[.8]	[1.1]	[.9]
Critical thinking	2.6	2.1	2.3	2.1	2.5	2.2	3.1	3.4	2.2	2.5	2.3	2.4	3.1	3.7
	(0)	(.1)	(.1)	(.1)	(.1)	(.1)	(.2)	(.1)	(.1)	(.1)	(.1)	(.1)	(.2)	(.3)
	[1]	[.7]	[.6]	[.7]	[.6]	[.9]	[1.2]	[.8]	[.9]	[.9]	[1]	[.9]	[1.2]	[1.1]
Autonomy	2.7	2.4	2.7	2.4	2.3	2.7	3.1	3.1	2.5	2.8	2.6	2.5	2.8	3.2
	(0)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(0)	(.1)	(.2)
	[.9]	[.7]	[.7]	[.6]	[.6]	[1]	[1]	[.9]	[.9]	[.9]	[1]	[.7]	[.7]	[.7]
Perseverance	2.3	2	2	1.9	2.2	1.7	2.8	2.9	2.1	2.4	2.3	2.3	2.3	2.7
	(0)	(.1)	(0)	(0)	(.1)	(.1)	(.2)	(.1)	(.1)	(.1)	(.1)	(0)	(.1)	(.2)
	[.8]	[.6]	[.5]	[.2]	[.5]	[.8]	[1.3]	[.8]	[.8]	[.8]	[1]	[.6]	[.6]	[.7]
Social and collaborative skills	2	1.5	1.2	2.1	1.5	1.1	3	2	2	2.1	2.2	2	2.5	2.7
	(0)	(.1)	(.1)	(.1)	(.1)	(0)	(.2)	(.1)	(.1)	(.1)	(.1)	(.1)	(.2)	(.3)
	[1]	[1.5]	[1.2]	[2.1]	[1.5]	[1.1]	[3]	[2]	[2]	[2.1]	[2.2]	[2]	[2.5]	[2.7]

Table shows Teach scores and scores for each sub construct in the Teach instrument, with standard errors in parentheses and standard deviation in brackets. Teach score measures the average across all items in the Teach instrument. Scores in the overall column are averages of scores across countries. Values in country columns are weighted.

Table A5: Teacher Content Knowledge

		Low-income						Lower-middle income					Upper-middle income	
	All	Ethiopia	Niger	Rwanda	Sierra Leone	Chad	Jordan	Pakistan ICT	Pakistan KP	Pakistan Balochistan	Pakistan Sindh	Pakistan Punjab	Peru	Gabon
Number of teachers (content)	2144	236	246	107	195	220	153	58	163	148	157	179	158	124
Content knowledge	53.9	53	48.9	61.9	37.9	46.7	63.7	43.3	48.2	58.3	52.2	69.8	67.5	52.4
	(.9)	(2.5)	(3.1)	(3.5)	(2.5)	(2.5)	(5)	(4.3)	(1.7)	(2.3)	(2.2)	(1.6)	(5.6)	(3.3)
	[22.3]	[19.4]	[21.3]	[21.1]	[17.7]	[20.9]	[22.4]	[19.7]	[16.7]	[22.7]	[22.5]	[18.2]	[23.6]	[18.3]
Number of teachers (literacy)	1210	103	147	67	97	107	86	28	92	104	125	137	66	51
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]
Literacy content knowledge	53.6	42.1	45.4	55.5	39.6	49.6	53.9	49.7	53.6	62.4	53.1	67.2	68.9	52.8
	(1.1)	(2)	(3.9)	(4.2)	(3.2)	(3.7)	(5.5)	(7.1)	(1.8)	(2.7)	(2.4)	(1.9)	(3.9)	(2.2)
	[20.9]	[2]	[3.9]	[4.2]	[3.2]	[3.7]	[5.5]	[7.1]	[1.8]	[2.7]	[2.4]	[1.9]	[3.9]	[2.2]
Number of teachers (mathematics)	936	133	99	40	98	113	67	31	71	45	32	42	92	73
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]
Mathematics content knowledge	54.2	63.4	53.8	75.1	35.1	44.3	84.5	31.7	40.7	48.2	47.7	79.7	66.5	52.1
	(1.7)	(2.3)	(4)	(4.1)	(3.8)	(3.4)	(1.7)	(1.6)	(2.8)	(2.9)	(4.6)	(1.8)	(9)	(4.9)
	[24.3]	[2.3]	[4]	[4.1]	[3.8]	[3.4]	[1.7]	[1.6]	[2.8]	[2.9]	[4.6]	[1.8]	[9]	[4.9]

Sample is restricted to teachers selected for the classroom observation module who also completed the content knowledge assessment. Table shows teacher content knowledge scores with standard error in parentheses and standard deviation in brackets. Teacher content knowledge measures the percentage of correct questions on the teacher content knowledge test (language for language teachers and mathematics for mathematics teachers), ranging from 0 to 100. Scores in the overall column are averages of scores across countries. Values in country columns are weighted.

Table A6: Policy Levers for Teaching Quality

		Low income					Lower-middle income						Upper-middle income	
	All	Ethiopia	Niger	Rwanda	Sierra Leone	Chad	Jordan	Pakistan ICT	Pakistan KP	Pakistan Balochistan	Pakistan Sindh	Pakistan Punjab	Peru	Gabon
Recruitment score (0-9)	2.6	3.4	2.2	4.2	2.2	2.5	2.3	2.4	2.8	2.2	2	2.9	2.3	2.2
	(.1)	(.2)	(.1)	(.2)	(.2)	(.2)	(.4)	(.1)	(.1)	(.1)	(.1)	(.1)	(.3)	(.3)
	[1.5]	[1.7]	[1.6]	[1.9]	[1.3]	[1.4]	[2.1]	[.9]	[1]	[1.1]	[1.1]	[1.1]	[1.4]	[1.4]
Completed required coursework	25.9	52.6	8.5	36.9	20.3	25	22.5	13.2	24.8	26.5	36.7	33.3	15.9	8.3
	(1.6)	(7.1)	(2.7)	(5.6)	(5.3)	(5)	(4.9)	(4.9)	(4.4)	(3.6)	(4.9)	(3.3)	(6.3)	(4.3)
	[43.8]	[50.1]	[28]	[48.6]	[40.5]	[43.4]	[42]	[34.2]	[43.3]	[44.3]	[48.3]	[47.3]	[36.7]	[27.8]
Achieved special educational qualification	59.2	62.5	70.8	57.5	78.5	45.7	85.5	33	50.2	47	55.1	71	37.1	78.2
	(2.1)	(6.1)	(4.8)	(6.2)	(6)	(6.1)	(5.9)	(9.1)	(5.3)	(5.1)	(4.5)	(4)	(9.2)	(11.4)
	[49.2]	[48.5]	[45.6]	[49.8]	[41.4]	[50]	[35.4]	[47.5]	[50.2]	[50.1]	[49.9]	[45.5]	[48.5]	[41.6]
Graduated from tertiary education degree program	25.3	11.2	15.7	85.8	24.6	23.6	15.9	30.3	21	19.5	22	19.9	12.1	20.8
	(1.8)	(3.3)	(4.5)	(4.1)	(7.5)	(5.3)	(5.1)	(9.1)	(4.1)	(3.6)	(3.8)	(3.5)	(7.2)	(11.4)
	[43.5]	[31.6]	[36.5]	[35.1]	[43.3]	[42.6]	[36.8]	[46.4]	[40.9]	[39.7]	[41.6]	[40]	[32.8]	[40.9]
Graduated from tertiary program designed to prepare teachers	27.9	24.8	30.6	35.1	27.8	39.1	13.1	68.3	23.1	23.7	12.5	18.2	30.8	27.8
	(2)	(4.7)	(4.8)	(6.2)	(6.6)	(6.5)	(5.3)	(8.3)	(4.4)	(4.3)	(2.2)	(2.7)	(10)	(12.9)
	[44.9]	[43.3]	[46.2]	[48.1]	[45.1]	[49]	[33.9]	[47]	[42.3]	[42.7]	[33.2]	[38.7]	[46.4]	[45.1]
Passed a subject content knowledge written test	42.4	61.3	36	78.1	14.6	40.6	15	38.7	73.4	36.9	24.1	45.4	43	28.9
	(2.1)	(7.4)	(5.7)	(5.1)	(4.9)	(6.2)	(5.3)	(9.3)	(4.4)	(4.3)	(4.2)	(3.8)	(9.4)	(15.5)
	[49.4]	[48.9]	[48.2]	[41.6]	[35.6]	[49.3]	[35.9]	[49.2]	[44.3]	[48.4]	[42.9]	[49.9]	[49.7]	[45.7]
Passed an interview stage assessment	41.8	66.9	9.8	49	16.5	36.6	30.8	52.5	62	56.2	45.8	72	22.1	17.1
	(2)	(7.2)	(3)	(6.1)	(4.9)	(5.9)	(7.4)	(10.1)	(5.1)	(4.9)	(4.3)	(3.4)	(9.5)	(7.4)
	[49.3]	[47.2]	[29.8]	[50.4]	[37.4]	[48.3]	[46.4]	[50.4]	[48.7]	[49.8]	[50]	[45]	[41.6]	[38]
Had a minimum amount of practical professional experience	17.5	17.7	17.8	11	28.4	20.1	17.3	4.6	23	9.8	6.7	23.7	11.8	33.3
	(1.8)	(5.1)	(4.2)	(3.7)	(7.5)	(4.4)	(5.5)	(2.6)	(4.9)	(3.3)	(2.4)	(3.8)	(4.2)	(15.6)
	[38.1]	[38.3]	[38.4]	[31.5]	[45.4]	[40.2]	[38]	[21.2]	[42.2]	[29.9]	[25.1]	[42.6]	[32.4]	[47.5]
Passed a practical assessment conducted by a supervisor	14.4	24.3	22.2	37.6	7.2	12.9	20.3	2.6	2	4.4	1.8	1.7	40.3	5.1
	(1.5)	(5.1)	(5.2)	(5.9)	(3.4)	(4.5)	(7)	(2.5)	(1.4)	(1.9)	(1.5)	(1)	(9.6)	(3.2)
	[35.1]	[43]	[41.7]	[48.8]	[26]	[33.7]	[40.5]	[16]	[14.1]	[20.5]	[13.3]	[13]	[49.3]	[22.2]
Conduct during mockup class	6.6	14.8	5.9	27.6	4.4	1.8	10.7	1.2	.1	0	0	1.4	14.4	.7
	(1.1)	(5.1)	(2)	(5.4)	(3)	(1.8)	(5)	(1.2)	(.1)	(.)	(.)	(1.1)	(8.7)	(.5)
	[24.8]	[35.6]	[23.6]	[45]	[20.5]	[13.4]	[31.1]	[10.9]	[3.2]	[0]	[0]	[11.9]	[35.3]	[8.2]
Probationary period for new teachers	46.4	47.6	40	64.5	49.7	36.9	48.4	81.1	80.7	22.7	12	47.5	13.5	79.7
	(2.2)	(6.2)	(6.1)	(5.9)	(8.2)	(5.9)	(9.6)	(11.5)	(4.3)	(3.5)	(3.2)	(4.3)	(4.7)	(11.2)
	[49.9]	[50.1]	[49.1]	[48.2]	[50.3]	[48.4]	[50.3]	[39.5]	[39.6]	[42]	[32.6]	[50.1]	[34.3]	[40.5]

Participated in induction/mentorship program	44.7	55	50.9	48.8	58.1	39.6	50	37.6	41.2	9.6	28.5	56.7	39.7	64.9
	(2.2)	(6.4)	(5.9)	(6.3)	(8.4)	(6.1)	(10.1)	(11.1)	(5.4)	(2.8)	(4.2)	(4)	(10.1)	(15.3)
	[49.7]	[49.9]	[50.1]	[50.4]	[49.7]	[49.1]	[50.3]	[48.9]	[49.4]	[29.5]	[45.3]	[49.7]	[49.1]	[48.1]
Teaching practicum under pre-service training	33.4	37.1	52.2	22.5	62.5	38.5	19.5	17.4	6.6	8.2	21.9	24.9	63.4	51.8
	(2)	(6.3)	(6)	(4.7)	(8.4)	(6.4)	(7.7)	(7.9)	(2.1)	(3)	(3.9)	(3.5)	(9.5)	(16.3)
	[47.2]	[48.4]	[50.1]	[42]	[48.7]	[48.8]	[39.9]	[38.3]	[24.8]	[27.5]	[41.5]	[43.4]	[48.4]	[50.3]
Hours per day spent teaching during teaching practicum	1.3	.9	3.1	1.1	2	1.5	.6	.1	.2	.2	.9	.4	3	3.1
	(.1)	(.2)	(.4)	(.3)	(.4)	(.3)	(.2)	(.1)	(.1)	(.1)	(.2)	(.1)	(.5)	(1.1)
	[2.3]	[1.4]	[3.1]	[2.6]	[2.3]	[2.1]	[1.3]	[.4]	[.7]	[.7]	[1.9]	[1]	[2.6]	[3.4]
Participated in in-service training	39.2	16.9	33.2	59.9	43.4	40.9	18.7	69.2	48.8	5.6	9.6	36.4	92.7	35.9
	(2.2)	(5.2)	(5.6)	(5.9)	(8.5)	(6.4)	(5.8)	(8.2)	(5.2)	(2.7)	(2.7)	(4)	(3.1)	(16.7)
	[48.8]	[37.6]	[47.3]	[49.4]	[49.9]	[49.3]	[39.2]	[46.6]	[50.1]	[23.1]	[29.5]	[48.2]	[26]	[48.3]
Days in-service training lasted	3.3	1.6	3.1	3.8	1.7	5.1	2.9	3.7	2.6	1.9	1.3	2.9	10.4	.9
	(.4)	(.8)	(.9)	(.9)	(.3)	(2)	(1.1)	(1.6)	(.4)	(1.5)	(.4)	(.5)	(3.5)	(.4)
	[9]	[4.6]	[7.7]	[7.2]	[2.4]	[14]	[8.6]	[5]	[3.6]	[11.5]	[5]	[8.1]	[16.3]	[2.4]
No in service training done in classroom	84.8	84.2	85.2	82.8	95.1	60.1	84.1	92.7	96.4	100	98	98.6	40.1	93.1
	(1.6)	(5.2)	(4.3)	(4.7)	(2.2)	(6.3)	(5.5)	(3.6)	(1.4)	(.)	(1)	(1)	(8.8)	(3.9)
	[35.9]	[36.6]	[35.7]	[38]	[21.7]	[49.1]	[36.8]	[26.3]	[18.6]	[0]	[14.2]	[12]	[49.2]	[25.5]
Less than 50% in service training done in classroom	9.2	12.2	10	15.5	2.4	28.7	1.8	7.3	3	0	1.6	1.4	22.3	6.3
	(1.2)	(4.9)	(4.1)	(4.7)	(1.4)	(6.1)	(1.1)	(3.6)	(1.4)	(.)	(.9)	(1)	(7.7)	(3.8)
	[28.9]	[32.9]	[30.1]	[36.4]	[15.3]	[45.4]	[13.2]	[26.3]	[17.2]	[0]	[12.5]	[12]	[41.8]	[24.5]
Over 50% in service training done in classroom	6.1	3.6	4.8	1.7	2.5	11.2	14.1	0	.6	0	.4	0	37.6	.6
	(1.2)	(1.7)	(1.9)	(.2)	(1.6)	(3)	(5.3)	(.)	(.4)	(.)	(.5)	(.)	(10.1)	(.4)
	[23.9]	[18.7]	[21.5]	[13.1]	[15.8]	[31.6]	[35]	[0]	[7.5]	[0]	[6.7]	[0]	[48.6]	[7.6]
Classroom has been observed	74.9	83	65.8	89.4	94.1	73.9	62.8	96.2	72.8	45.4	36	76.8	93.9	91.1
	(1.6)	(4.6)	(5.7)	(4.3)	(2.8)	(5.3)	(10.9)	(2.8)	(4.9)	(4.8)	(4)	(3.2)	(3.9)	(4.5)
	[43.4]	[37.7]	[47.6]	[31]	[23.8]	[44]	[48.6]	[19.3]	[44.6]	[50]	[48.1]	[42.3]	[24]	[28.6]
Months since last classroom observation	1.9	3.4	1.8	3	2.3	2	1	1.1	.7	3.2	1.2	1	2.2	1.9
	(.2)	(.8)	(.3)	(.3)	(.3)	(.2)	(.4)	(.3)	(.1)	(2)	(.2)	(.1)	(.4)	(.5)
	[5]	[4.2]	[4]	[2.5]	[3.6]	[2.6]	[2.3]	[1.6]	[1]	[15.4]	[2.6]	[1.5]	[2.5]	[2.2]
Discussed results of observation	58.4	78.4	42.8	78.7	72.8	56.5	46.1	75.3	49.5	17.8	29.3	57	76.4	83.2
	(2)	(5.1)	(5.3)	(5.4)	(6.6)	(6.3)	(8.9)	(7.9)	(5.2)	(3.4)	(3.7)	(4.3)	(8)	(7.3)
	[49.3]	[41.3]	[49.6]	[41.2]	[44.8]	[49.7]	[50.1]	[43.6]	[50.2]	[38.4]	[45.7]	[49.6]	[42.6]	[37.6]
Observer provided feedback	53.8	75.2	39.7	70.8	67.3	49.5	41.9	70.9	48.1	12.4	23.4	52.2	73.5	79.7
	(2.1)	(5.3)	(5.3)	(5.7)	(7.4)	(6.4)	(8.5)	(8.6)	(5.2)	(2.7)	(3.5)	(4.2)	(8.3)	(8.4)
	[49.9]	[43.3]	[49.1]	[45.8]	[47.2]	[50.2]	[49.6]	[45.9]	[50.1]	[33.1]	[42.5]	[50.1]	[44.3]	[40.5]
Had lesson plans for last week	73.7	97.6	51.7	77.8	60.6	78.7	96.7	92.7	61.9	38.3	59.8	67.8	99.3	89.2
	(1.7)	(.9)	(5.8)	(6.2)	(8.8)	(5)	(1.5)	(3.1)	(5.3)	(4.6)	(4.4)	(4.1)	(.4)	(6.1)

	[44]	[15.4]	[50.1]	[41.9]	[49.2]	[41.1]	[17.9]	[26.2]	[48.7]	[48.8]	[49.2]	[46.9]	[8.6]	[31.3]
Discussed lesson plans	44.5	77.5	17	38.3	56	37.7	56.9	67.8	42.7	16.5	20	27.1	86.8	48.6
	(2.1)	(6.1)	(4.8)	(6)	(8.6)	(5.9)	(10.2)	(9.9)	(5.3)	(3.1)	(3)	(3.6)	(5.9)	(16.1)
	[49.7]	[41.9]	[37.7]	[49]	[49.9]	[48.6]	[49.8]	[47.2]	[49.6]	[37.3]	[40.1]	[44.6]	[34]	[50.3]
Evaluated on pedagogical skill and content knowledge	53.8	73.9	44.5	89.6	61.7	57.9	35.9	40.7	42.9	20.1	20.6	64.3	87.8	40.5
	(2.1)	(7)	(5.8)	(4.4)	(8.1)	(5.9)	(7.8)	(11.9)	(5.3)	(3.7)	(3.2)	(3.8)	(3.4)	(15.4)
	[49.9]	[44]	[49.9]	[30.8]	[48.9]	[49.5]	[48.2]	[49.6]	[49.6]	[40.2]	[40.6]	[48.1]	[32.8]	[49.5]
Financial consequences to negative evaluation	81.6	89.2	93.6	77	93.8	70.2	63.2	73.8	63.7	72.4	87.2	86.3	99.2	79.7
	(1.6)	(2.3)	(2.4)	(5.8)	(2.4)	(6)	(8.9)	(9.5)	(5.2)	(4.7)	(2.5)	(3.5)	(.4)	(12.5)
	[38.8]	[31.2]	[24.5]	[42.4]	[24.3]	[45.9]	[48.5]	[44.4]	[48.2]	[44.9]	[33.5]	[34.5]	[9.2]	[40.5]
Financial consequences to positive evaluation	70.2	56.9	78.1	30.2	72	60.8	55.6	62.3	73.5	82.2	91.2	94.3	70.6	77.5
	(2)	(6.2)	(5.3)	(5.5)	(7.1)	(5.9)	(10.3)	(11.3)	(4.7)	(3.7)	(3.1)	(1.5)	(8.8)	(12.6)
	[45.8]	[49.7]	[41.5]	[46.3]	[45.2]	[49]	[50]	[48.9]	[44.3]	[38.4]	[28.4]	[23.3]	[45.7]	[42.1]
Meritocracy for advancement	50.4	43.8	60.9	87.3	94.2	61.1	42.5	65.8	33.8	36.5	17.7	18.2	49.6	50.9
	(2.2)	(6.3)	(5.9)	(4)	(2.2)	(5.9)	(9.8)	(7.1)	(4.9)	(4.7)	(3.1)	(2.9)	(9.9)	(15.5)
	[50]	[49.8]	[48.9]	[33.5]	[23.5]	[48.9]	[49.7]	[47.9]	[47.4]	[48.3]	[38.3]	[38.7]	[50.2]	[50.4]
Received bonus during last academic year	9.9	7	12	42.1	3.2	1.5	27.4	1.7	1.1	2.2	0	0	8	27.2
	(1.7)	(3.2)	(4.7)	(6.3)	(2)	(.7)	(7.2)	(.9)	(.8)	(.9)	(.)	(.)	(4.1)	(15.7)
	[29.9]	[25.6]	[32.6]	[49.7]	[17.6]	[12.3]	[44.8]	[12.9]	[10.3]	[14.7]	[0]	[0]	[27.2]	[44.8]
Salary was delayed during last academic year	16.2	16	26.3	30.7	16.8	30.9	4.3	37	8.1	10.9	2.6	.9	6.4	25.1
	(1.6)	(6.3)	(4.8)	(6.2)	(5.5)	(5.3)	(1.8)	(10.8)	(3)	(3.7)	(1.5)	(.7)	(3.9)	(11.9)
	[36.8]	[36.7]	[44.2]	[46.5]	[37.6]	[46.3]	[20.3]	[48.8]	[27.4]	[31.3]	[16]	[9.3]	[24.6]	[43.6]
Motivation - always wanted to be a teacher and likes teaching	55.7	65.9	53.5	60.5	32.9	63.2	36.6	29.1	61	66	82.9	67.7	38.7	65.7
	(2)	(4.9)	(6.5)	(5.5)	(6.4)	(5.9)	(7.2)	(6.3)	(5.3)	(4.1)	(3.2)	(4.2)	(8.5)	(11.3)
	[49.7]	[47.5]	[50]	[49.2]	[47.2]	[48.4]	[48.3]	[45.7]	[48.9]	[47.5]	[37.7]	[46.9]	[48.9]	[47.7]
Motivation - teaching offers steady career path	7.4	4.9	2	7.5	16.8	5.5	14.6	6.5	6.1	4.8	5.8	16.5	.6	4.4
	(.9)	(2.1)	(1.2)	(3)	(5.7)	(2.2)	(4.6)	(3.7)	(2.1)	(1.9)	(2.2)	(3.4)	(.4)	(2.7)
	[26.2]	[21.6]	[13.9]	[26.6]	[37.6]	[22.8]	[35.5]	[24.8]	[24.1]	[21.4]	[23.5]	[37.3]	[7.9]	[20.6]
Motivation - benefit the socially disadvantaged	17.8	9.7	24.1	24.4	32.9	24.9	2.2	14.3	16.3	6.5	8.7	8.3	49.9	8.6
	(1.6)	(2.3)	(5.1)	(5.1)	(7.8)	(5.9)	(1.6)	(7.2)	(3.7)	(2.3)	(2.1)	(2.4)	(8.9)	(4.6)
	[38.2]	[29.6]	[42.9]	[43.3]	[47.2]	[43.4]	[14.8]	[35.2]	[37.1]	[24.8]	[28.2]	[27.7]	[50.2]	[28.2]
Intrinsic motivation	3.8	3.7	3.6	4	3.6	3.6	3.5	4	4	3.6	3.4	3.7	4.1	4.2
	(0)	(.1)	(.1)	(0)	(.1)	(0)	(.1)	(0)	(.1)	(0)	(0)	(0)	(0)	(.1)
	[.5]	[.4]	[.5]	[.3]	[.6]	[.4]	[.5]	[.3]	[.5]	[.5]	[.3]	[.5]	[.3]	[.3]
Teacher monitoring	2.6	3.2	1.8	3	2.3	2.5	3	2.2	2.7	2.6	2.5	2.6	2.8	2.1
	(0)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.2)
	[.7]	[.6]	[.6]	[.5]	[.5]	[.6]	[.6]	[.5]	[.7]	[.8]	[.6]	[.7]	[.4]	[.7]
Teaching evaluation	3.4	4.3	3.3	4.6	3.8	3.6	3.4	3	2.9	2.4	2.1	3.3	4.1	2.8
	(0)	(.1)	(.1)	(0)	(.1)	(.1)	(.1)	(.2)	(.1)	(.1)	(.1)	(.1)	(.1)	(.3)
	[1.1]	[.5]	[.7]	[.3]	[.7]	[.8]	[.8]	[1]	[.9]	[1.1]	[1.1]	[.9]	[.5]	[.9]
Teacher support	2.6	2.7	3	3	3.2	2.9	2.4	2.6	2.5	1.5	1.9	2.6	3.2	3
	(0)	(.1)	(.2)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)

	[.9]	[.9]	[.9]	[.4]	[.7]	[.8]	[.9]	[.6]	[.7]	[.7]	[1]	[.8]	[.6]	[.5]
Teacher attraction	3.5	3.5	3.6	3.9	3.7	3.4	3.6	3.5	3.5	3.5	3.3	3.3	3.5	3.3
	(0)	(0)	(.1)	(0)	(0)	(.1)	(.1)	(0)	(0)	(0)	(0)	(0)	(.1)	(.1)
	[.4]	[.4]	[.5]	[.3]	[.4]	[.5]	[.5]	[.3]	[.4]	[.5]	[.4]	[.3]	[.3]	[.4]
Instructional leadership	3.2	3.9	2.6	3.5	3.6	3.1	3.2	3.4	2.9	2.2	2.3	2.9	4	3.5
	(0)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.1)	(.2)
	[1]	[.8]	[1.1]	[.5]	[.8]	[.9]	[.8]	[.8]	[.9]	[.9]	[1.1]	[.9]	[.6]	[.8]

Sample is restricted to teachers selected for the classroom observation module. Table shows mean with standard error in parentheses and standard deviation in brackets. A recruitment score is between 0 and 9. It measures the number of factors teacher selected under the factors considered during the recruitment process of new teachers. Financial consequences includes if teacher is dismissed/promoted and if teacher's salary would be reduced/increased. All school level indicators are average scores between 1 (worst) and 5 (best) at the school level.

Table A7: Results for Random Forest and Conditional Inference Forest Variable Selection Models

	Student Learning Regression		Teach Score Regression	
	Importance Rank: Random Forest Model	Importance Rank: Conditional Inference Forest Model	Importance Rank: Random Forest Model	Importance Rank: Conditional Inference Forest Model
Student Attendance	11	28	4	21
Time (in minutes) spent on reading practice	13	71	6	75
Operational management	16	17	23	11
Availability of textbooks	125	4	183	3
Average 1st-grade student knowledge score	19	1	5	5

Annex B – Detail on control variables

Below, we provide details on independent variables included in the regressions to explain teacher pedagogical skill and content knowledge:

- C_{ijk} - is a vector for recruitment and training indicators. We consider what factors were considered when recruiting new teachers, whether the teacher went through a probationary period as a new teacher, and whether the teacher participated in an induction program. For pre-service and in-service training, we consider whether the teacher was required to attend a teaching practicum during pre-service and the number of hours spent teaching during the practicum. We also consider if the teacher received in-service training, how long the training was for (in days) and if there was an in-classroom component.

- D_{ijk} - is a vector for instructional leadership indicators. We consider if a classroom has been observed before, when was the last observation, was feedback provided, and if the teacher had lesson plans and discussed the plan with someone else before teaching it.

- E_{ijk} - is a vector for evaluation indicators. We consider whether a teacher was evaluated on pedagogical skill or content knowledge, and consequences when a teacher receives a positive evaluation and a negative evaluation.

- F_{ijk} – is a vector for incentives indicators. We consider whether meritocracy is recognized, if bonuses were received during the last academic year, and if salary was delayed during the last academic year.

- G_{ijk} – is a vector for intrinsic motivation indicators. We consider if the respondent chose to become a teacher because they aspired to be a teacher and liked the profession or believe that teaching provides a steady career path, or believe that teaching can help the socially disadvantaged and shape child and adolescent values.

- H_{jk} - is a vector of school level indicators including teacher support, school support, teaching evaluation, instructional leadership, intrinsic motivation and teacher attraction. All school level indicators are average

scores between 1 (worst) and 5 (best) at the school level. Intrinsic motivation measures whether teachers are intrinsically motivated to teach. The question(s) aim to address this phenomenon by measuring the level of intrinsic motivation among teachers as well as teacher values that may be relevant for ensuring that the teacher is motivated to focus on all children and not just some. It measures teacher's perception on if it is acceptable for a teacher to be absent (if the assigned curriculum has been completed, if students are left with work to do, if the teacher is doing something useful for the community), if students deserve more attention than others (if they attend school regularly, if they come to school with materials, and if they are motivated to learn), if students can change their level of intelligence, and the reason why each teacher is motivated to teach. Teacher monitoring measures the extent to which teacher presence is being monitored, whether attendance is rewarded, and whether there are consequences for chronic absence. Score is the sum of the following: 1 point if teachers were evaluated by some authority on basis of absence, 1 point if good attendance is rewarded, 1 point if there are consequences for chronic absence (more than 30% absence) and 1 point minus the fraction of teachers that had to miss class because of any of the following: collect paycheck, school administrative procedure, errands or request of the school district office, other administrative tasks. Teaching evaluation measures whether there is a teacher evaluation system in place, and if so, the types of decisions that are made based on the evaluation results. Score is the sum of the following: 1 point if teacher formally evaluated in past school year, 1 point (0.2 points for each of the following: evaluation included evaluation of attendance, knowledge of subject matter, pedagogical skills in the classroom, students' academic achievement, students' socio-emotional development), 1 point if consequences exist if teacher receives 2 or more negative evaluations and 1 point if rewards exist if teacher receives 2 or more positive evaluations. Teacher support asks teachers about participation and the experience with several types of formal/informal training. The score includes: 1 point for pre-service training (0.5 points if teacher had pre-service training and 0.5 points if teacher reported receiving usable skills from training), 1 point for teaching practicum (0.5 points if teacher participated in a practicum and 0.5 points if practicum lasted more than 3 months and teacher spent more than one hour per day teaching to students), 1 point for in-service training (0.5 points if teacher had in-service training, 0.25 points if in-service training lasted more than 2 total days, 0.125 points if more than 25% of the in-service training was done in the classroom and 0.125 points if more than 50% of the in-service training was done in the classroom), and 1 point if opportunities for teachers to come together to share ways of improving teaching exist. For the teacher attraction score, 0.8 points is awarded for each of the following: teacher is satisfied with job, teacher is satisfied with status in community, would better teachers be promoted faster?, do teachers receive bonuses?, one minus the fraction of months in past year with a salary delay. Instructional leadership score starts at 1 and points added are the sum of whether a teacher: had a classroom observation in past year, had a discussion based on that observation that lasted longer than 10 min, received actionable feedback from that observation, and had a lesson plan and discussed it with another person.

Annex C – Additional Sensitivity Checks

Table C1: Regression Results for Core Model (By Income Group)

(a) Full Sample

	(1) Student knowledge (Low income)	(2) Literacy student knowledge (Low income)	(3) Math student knowledge (Low income)	(4) Student knowledge (Lower-middle income)	(5) Literacy student knowledge (Lower-middle income)	(6) Math student knowledge (Lower-middle income)	(7) Student knowledge (Upper-middle income)	(8) Literacy student knowledge (Upper-middle income)	(9) Math student knowledge (Upper-middle income)
Panel A: Teacher pedagogical skill									
Teach score	1.21	0.61	1.82	2.45***	3.13***	1.77*	3.81*	5.36***	2.26
	(1.25)	(1.41)	(1.36)	(0.80)	(0.87)	(0.91)	(2.08)	(1.93)	(2.53)
Observations	20916	20916	20916	13828	13828	13828	4168	4168	4168
R^2	0.179	0.243	0.083	0.166	0.217	0.078	0.227	0.220	0.131
Panel B: Teacher content knowledge									
Content knowledge	0.088***	0.078***	0.099***	0.036	0.029	0.042	0.13***	0.052	0.20***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)
Observations	18691	18691	18691	13449	13449	13449	4058	4058	4058
R^2	0.189	0.262	0.083	0.161	0.216	0.074	0.239	0.199	0.165
Panel C: Pedagogical skill and content knowledge									
Teach score	0.61	0.13	1.10	2.34***	2.80***	1.87**	3.46	5.34**	1.58
	(1.29)	(1.43)	(1.43)	(0.81)	(0.85)	(0.92)	(2.18)	(2.06)	(2.61)
Content knowledge	0.087***	0.078***	0.096***	0.034	0.027	0.041	0.13***	0.051	0.20***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)
Observations	18691	18691	18691	13449	13449	13449	3939	3939	3939
R^2	0.189	0.262	0.083	0.166	0.221	0.077	0.240	0.215	0.160
Panel D: Pedagogical skill sub-scores									
Classroom culture	0.78	0.37	1.20	0.88	1.37	0.40	1.05	1.90	0.20
	(1.02)	(1.20)	(1.13)	(0.91)	(1.02)	(1.00)	(2.36)	(2.07)	(3.17)
Instruction	-0.89	-0.33	-1.45	1.80*	1.22	2.38**	0.86	1.66	0.056
	(1.03)	(1.21)	(1.24)	(0.96)	(1.06)	(1.09)	(1.73)	(1.56)	(2.48)
Socio emotional skills	1.52	0.65	2.38*	-0.21	0.57	-0.98	1.96	1.86	2.07
	(1.11)	(1.33)	(1.22)	(0.91)	(1.00)	(1.01)	(2.16)	(1.95)	(3.02)
Observations	20916	20916	20916	13828	13828	13828	4168	4168	4168
R^2	0.180	0.244	0.085	0.167	0.217	0.080	0.227	0.220	0.131
School characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Student characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Extended controls	No	No	No	No	No	No	No	No	No
Survey time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table presents point estimates from regressing student learning (student knowledge score) on dimensions of teaching quality using our core model for low-income (Columns 1-3), lower-middle-income (Columns 4-6), and upper-middle-income countries (Columns 7-9). Columns 1, 4, and 7 refer to overall student knowledge, columns 2, 5, and 8 refer to student knowledge in literacy and columns 3, 6, and 9 refer to student knowledge in mathematics. In panel A, teacher pedagogical skill is measured by the TEACH score (1-5), an average of all items in the TEACH instrument. In panel B, teacher content knowledge measures the percentage of correct questions on the teacher content knowledge test (language for language teachers and mathematics for mathematics teachers), ranging from 0 to 100. Panel C includes both pedagogical skill and content knowledge, as described earlier, while panel D includes pedagogical skills sub-scores. In all columns, the model controls for a vector of student, teacher, and school characteristics. Student controls include age and gender.. Teacher controls include gender, teaching experience in years, and teacher's level of education. School controls include rurality, number of students enrolled, average teacher absence rate, infrastructure index, and a school-level wealth index. The infrastructure index measures whether a school has access to drinking water, functioning toilet, functioning internet, and functioning electricity in the classroom, and whether the school is accessible to students with disabilities. These models also control for survey time, the number of months from the start of the academic year to implementation of survey instrument. All models include country fixed effects. Standard errors are clustered by school and reported in parentheses. * p< 0.1, ** p< 0.05, *** p< 0.01

(b) Restricted Sample Based on Teacher Subject(s) Taught

	(1) Student knowledge (Low income)	(2) Literacy student knowledge (Low income)	(3) Math student knowledge (Low income)	(4) Student knowledge (Lower-middle income)	(5) Literacy student knowledge (Lower-middle income)	(6) Math student knowledge (Lower-middle income)	(7) Student knowledge (Upper-middle income)	(8) Literacy student knowledge (Upper-middle income)	(9) Math student knowledge (Upper-middle income)
Panel A: Teacher pedagogical skill									
Teach score	1.68	1.79	1.21	4.17***	4.23***	2.56**	3.81*	5.43***	2.12
	(1.62)	(1.54)	(1.56)	(1.16)	(1.02)	(1.03)	(2.13)	(1.97)	(2.58)
Observations	12372	17449	15814	8646	11835	10614	4101	4126	4115
R^2	0.180	0.236	0.081	0.181	0.208	0.104	0.227	0.216	0.128
Panel B: Teacher content knowledge									
Content knowledge	0.11***	0.099***	0.095***	0.074**	0.046	0.047	0.13***	0.053	0.21***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)
Observations	11947	15525	15088	8415	11510	10329	3966	3991	4005
R^2	0.197	0.259	0.086	0.174	0.209	0.097	0.235	0.191	0.163
Panel C: Pedagogical skill and content knowledge									
Teach score	1.09	1.31	0.74	3.84***	3.67***	2.73***	3.59	5.43**	1.66
	(1.59)	(1.54)	(1.57)	(1.16)	(0.98)	(1.04)	(2.23)	(2.11)	(2.67)
Content knowledge	0.11***	0.097***	0.093***	0.067**	0.044	0.047	0.13***	0.051	0.21***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)
Observations	11947	15525	15088	8415	11510	10329	3872	3897	3886
R^2	0.197	0.260	0.086	0.183	0.216	0.102	0.239	0.210	0.158
Panel D: Pedagogical skill sub-scores									
Classroom culture	2.56**	1.13	2.27*	0.081	1.25	-0.24	0.81	1.89	-0.12

	(1.25)	(1.31)	(1.31)	(1.16)	(1.11)	(1.14)	(2.37)	(2.07)	(3.18)
Instruction	-2.36*	0.18	-3.56**	2.67**	1.72	2.71**	0.90	1.67	0.13
	(1.33)	(1.29)	(1.45)	(1.21)	(1.11)	(1.25)	(1.74)	(1.56)	(2.49)
Socio emotional skills	1.57	0.47	2.80*	1.18	1.25	-0.018	2.11	1.92	2.10
	(1.46)	(1.43)	(1.47)	(1.21)	(1.13)	(1.11)	(2.18)	(1.97)	(3.04)
Observations	12372	17449	15814	8646	11835	10614	4101	4126	4115
R^2	0.185	0.236	0.087	0.182	0.208	0.105	0.227	0.216	0.128
School characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Extended controls	No	No	No	No	No	No	No	No	No
Survey time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table presents point estimates from regressing student learning (student knowledge score) on dimensions of teaching quality using our core model for low-income (Columns 1-3), lower-middle-income (Columns 4-6), and upper-middle-income countries (Columns 7-9). In this table, we restrict teachers depending on the subject taught: columns 1, 4, and 7 include only teachers who teach both language and math, columns 2, 5, and 8 include teachers who teach either both language and math or language only, and columns 3, 6, and 9 include teachers who teach either language and math or math only. Columns 1, 4, and 7 refer to overall student knowledge, columns 2, 5, and 8 refer to student knowledge in literacy and columns 3, 6, and 9 refer to student knowledge in mathematics. In panel A, teacher pedagogical skill is measured by the TEACH score (1-5), an average of all items in the TEACH instrument. In panel B, teacher content knowledge measures the percentage of correct questions on the teacher content knowledge test (language for language teachers and mathematics for mathematics teachers), ranging from 0 to 100. Panel C includes both pedagogical skill and content knowledge, as described earlier, while panel D includes pedagogical skills sub-scores. In all columns, the model controls for a vector of student, teacher, and school characteristics. Student controls include age and gender.. Teacher controls include gender, teaching experience in years, and teacher's level of education. School controls include rurality, number of students enrolled, average teacher absence rate, infrastructure index, and a school-level wealth index. The infrastructure index measures whether a school has access to drinking water, functioning toilet, functioning internet, and functioning electricity in the classroom, and whether the school is accessible to students with disabilities. These models also control for survey time, the number of months from the start of the academic year to implementation of survey instrument. All models include country fixed effects. Standard errors are clustered by school and reported in parentheses. * p< 0.1, ** p< 0.05, *** p< 0.01

Table C2: Regression Results for Core Model (By Country)

	(1) Ethiopia	(2) Niger	(3) Rwanda	(4) Sierra Leone	(5) Chad	(6) Jordan	(7) Pakistan ICT	(8) Pakistan KP	(9) Pakistan Balochistan	(10) Pakistan Sindh	(11) Pakistan Punjab	(12) Peru	(13) Gabon
Panel A: Teacher pedagogical skill													
Teach score	0.16	-1.29	1.81	2.63	1.17	0.13	1.62	5.86***	7.59**	-1.16	4.22**	1.87	6.40
	(2.24)	(3.09)	(3.22)	(3.20)	(2.10)	(1.05)	(1.61)	(2.07)	(3.57)	(2.50)	(1.70)	(1.28)	(4.17)
Observations	4461	4424	4431	3206	4394	2645	1410	3029	1422	1879	3443	2029	2139
R^2	0.087	0.211	0.163	0.212	0.119	0.062	0.140	0.176	0.135	0.079	0.154	0.202	0.125
Panel B: Teacher content knowledge													
Content knowledge	0.0045	0.046	0.052	-0.016	0.19***	-0.0011	-0.033	0.022	-0.030	-0.033	0.22***	0.12***	0.12**
	(0.05)	(0.05)	(0.06)	(0.05)	(0.04)	(0.04)	(0.05)	(0.05)	(0.10)	(0.10)	(0.06)	(0.04)	(0.06)
Observations	4359	4353	2541	3144	4294	2486	1410	3004	1422	1684	3443	1917	2141
R^2	0.089	0.213	0.174	0.209	0.170	0.051	0.139	0.150	0.104	0.073	0.175	0.224	0.126
Panel C: Pedagogical skill and content knowledge													

Teach score	0.50	-1.19	1.02	2.04	0.67	-0.18	1.81	5.90***	7.70**	-2.05	5.18***	0.90	6.87*
	(2.40)	(3.14)	(3.99)	(3.37)	(1.98)	(1.06)	(1.70)	(2.07)	(3.51)	(2.55)	(1.62)	(1.41)	(4.11)
Content knowledge	0.0010	0.047	0.047	-0.017	0.19***	-0.00052	-0.040	0.016	-0.045	-0.030	0.24***	0.11***	0.13**
	(0.05)	(0.04)	(0.06)	(0.05)	(0.04)	(0.04)	(0.05)	(0.05)	(0.11)	(0.10)	(0.06)	(0.04)	(0.06)
Observations	4359	4353	2541	3144	4294	2486	1410	3004	1422	1684	3443	1917	2022
R^2	0.089	0.213	0.175	0.210	0.171	0.051	0.141	0.170	0.136	0.076	0.189	0.225	0.144
Panel D: Pedagogical skill sub-scores													
Classroom culture	-4.08*	0.52	1.20	-1.32	4.40**	1.54	-1.33	0.027	2.75	-4.98	4.27**	0.61	-0.26
	(2.24)	(2.14)	(2.01)	(1.98)	(2.04)	(1.73)	(1.52)	(1.87)	(2.92)	(3.10)	(1.85)	(2.04)	(4.76)
Instruction	1.81	-5.70***	-0.64	3.17	-3.19	-1.44	3.13	6.57***	7.00**	6.48**	-3.08*	1.63	1.01
	(2.25)	(2.17)	(1.96)	(3.13)	(2.00)	(1.72)	(1.99)	(2.32)	(2.82)	(2.90)	(1.81)	(1.55)	(3.07)
Socio emotional skills	2.09	6.54**	1.40	1.42	-0.50	-0.000077	-0.30	-1.48	-2.12	-3.67	3.48	-0.72	5.75
	(2.31)	(2.90)	(2.21)	(2.94)	(2.10)	(1.41)	(1.88)	(2.17)	(3.90)	(2.48)	(2.25)	(2.03)	(4.18)
Observations	4461	4424	4431	3206	4394	2645	1410	3029	1422	1879	3443	2029	2139
R^2	0.096	0.226	0.164	0.215	0.135	0.064	0.145	0.190	0.151	0.104	0.165	0.203	0.129
School characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table presents point estimates from regressing student learning (student knowledge score) on dimensions of teaching quality using our core model for all sampled education systems. All columns report the effect on overall student knowledge. In panel A, teacher pedagogical skill is measured by the TEACH score (1-5), an average of all items in the TEACH instrument. In panel B, teacher content knowledge measures the percentage of correct questions on the teacher content knowledge test (language for language teachers and mathematics for mathematics teachers), ranging from 0 to 100. Panel C includes both pedagogical skill and content knowledge, as described earlier, while panel D includes pedagogical skills sub-scores. In all columns, the model controls for a vector of student, teacher, and school characteristics. Student controls include age and gender. Teacher controls include gender, teaching experience in years, and teacher's level of education. School controls include rurality, number of students enrolled, average teacher absence rate, infrastructure index, and a school-level wealth index. The infrastructure index measures whether a school has access to drinking water, functioning toilet, functioning internet, and functioning electricity in the classroom, and whether the school is accessible to students with disabilities. These models also control for survey time, the number of months from the start of the academic year to implementation of survey instrument. All models include country fixed effects. Standard errors are clustered by school and reported in parentheses. * p< 0.1, ** p< 0.05, *** p< 0.01

Table C3: Regression Results Controlling for Average Teacher Content Knowledge at the School Level

	(1) Student knowledge	(2) Student knowledge	(3) Student knowledge	(4) Literacy student knowledge	(5) Literacy student knowledge	(6) Literacy student knowledge	(7) Math student knowledge	(8) Math student knowledge	(9) Math student knowledge
Panel A: Teacher pedagogical skill									
Teach score	2.62***	1.78***	0.99	3.00***	1.93***	1.01	2.23***	1.64**	0.97
	(0.63)	(0.68)	(0.66)	(0.68)	(0.74)	(0.73)	(0.70)	(0.77)	(0.74)
Observations	38498	35044	31310	38498	35044	31310	38498	35044	31310
R^2	0.228	0.273	0.315	0.303	0.344	0.369	0.097	0.127	0.164
Panel B: Teacher content knowledge									
Content knowledge	0.064***	0.065***	0.053***	0.049**	0.046**	0.034	0.080***	0.083***	0.072***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	38618	35163	31383	38618	35163	31383	38618	35163	31383

R^2	0.226	0.273	0.315	0.300	0.343	0.370	0.097	0.127	0.164
Panel C: Pedagogical skill and content knowledge									
Teach score	2.47***	1.63**	0.85	2.89***	1.82**	0.92	2.04***	1.44*	0.77
	(0.63)	(0.68)	(0.66)	(0.68)	(0.73)	(0.73)	(0.71)	(0.77)	(0.74)
Content knowledge	0.059***	0.062***	0.050***	0.043**	0.043**	0.031	0.076***	0.080***	0.070***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	38498	35044	31310	38498	35044	31310	38498	35044	31310
R^2	0.230	0.275	0.317	0.304	0.345	0.370	0.101	0.131	0.167
Panel D: Pedagogical skill sub-scores									
Classroom culture	0.69	0.79	0.0059	0.63	0.74	-0.17	0.75	0.84	0.18
	(0.68)	(0.71)	(0.70)	(0.77)	(0.79)	(0.79)	(0.76)	(0.80)	(0.80)
Instruction	0.67	-0.13	-0.28	1.17	0.019	-0.25	0.17	-0.28	-0.31
	(0.72)	(0.71)	(0.70)	(0.81)	(0.79)	(0.80)	(0.84)	(0.86)	(0.85)
Socio emotional skills	1.27*	1.19*	1.30*	1.18	1.23	1.46*	1.36*	1.16	1.14
	(0.73)	(0.71)	(0.71)	(0.82)	(0.81)	(0.82)	(0.80)	(0.81)	(0.81)
Observations	38498	35044	31310	38498	35044	31310	38498	35044	31310
R^2	0.228	0.273	0.316	0.303	0.344	0.370	0.098	0.127	0.164
School characteristics	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Teacher characteristics	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Student characteristics	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Extended controls	No	No	Yes	No	No	Yes	No	No	Yes
Average teacher content knowledge at school	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey time	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table presents point estimates from regressing student learning (student knowledge score) on dimensions of teaching quality using our core model. Columns 1 and 2 refer to overall student knowledge, columns 3 and 4 refer to student knowledge in literacy and columns 5 and 6 refer to student knowledge in mathematics. In panel A, teacher pedagogical skill is measured by the TEACH score (1-5), an average of all items in the TEACH instrument. In panel B, teacher content knowledge measures the percentage of correct questions on the teacher content knowledge test (language for language teachers and mathematics for mathematics teachers), ranging from 0 to 100. Panel C includes both pedagogical skill and content knowledge, as described earlier, while panel D includes pedagogical skills sub-scores. In columns 2, 4, and 6, the model controls for a vector of student, teacher, and school characteristics. Student controls include age and gender. Teacher controls include gender, teaching experience in years, and teacher's level of education. School controls include rurality, number of students enrolled, average teacher absence rate, infrastructure index, and school-level wealth index. The infrastructure index measures whether a school has access to drinking water, functioning toilet, functioning internet, and functioning electricity in the classroom, and whether the school is accessible to students with disabilities. These models also control for survey time, the number of months from the start of the academic year to implementation of survey instrument, and the average teacher content knowledge at school. All models include country fixed effects. Standard errors are clustered by school and reported in parentheses. * p< 0.1, ** p< 0.05, *** p< 0.01

Table C4: Regression Results for Core Model (Excluding Teacher Education Variables)

	(1)	(2)	(3)
	Student knowledge	Literacy student knowledge	Math student knowledge
Panel A: Teacher pedagogical skill			
Teach score	1.96***	2.17***	1.76**
	(0.67)	(0.73)	(0.75)
Observations	38912	38912	38912

R2	0.267	0.327	0.131
Panel B: Teacher content knowledge			
Content knowledge	0.086***	0.073***	0.099***
	(0.02)	(0.02)	(0.02)
Observations	36198	36198	36198
R2	0.263	0.335	0.122
Panel C: Pedagogical skill and content knowledge			
Teach score	1.63**	1.78**	1.47*
	(0.68)	(0.73)	(0.77)
Content knowledge	0.084***	0.071***	0.097***
	(0.02)	(0.02)	(0.02)
Observations	36079	36079	36079
R2	0.266	0.336	0.125
Panel D: Pedagogical skill sub-scores			
Classroom culture	0.82	0.84	0.79
	(0.69)	(0.78)	(0.76)
Instruction	0.34	0.46	0.22
	(0.68)	(0.77)	(0.81)
Socio emotional skills	0.84	0.89	0.79
	(0.69)	(0.80)	(0.78)
Observations	38912	38912	38912
R2	0.267	0.327	0.131
School characteristics	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes
Student characteristics	Yes	Yes	Yes
Survey time	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes

This table presents point estimates from regressing student learning (student knowledge score) on dimensions of teaching quality using our core model. Column 1 refers to overall student knowledge, column 2 refers to student knowledge in literacy and column 3 refers to student knowledge in mathematics. In panel A, teacher pedagogical skill is measured by the TEACH score (1-5), an average of all items in the TEACH instrument. In panel B, teacher content knowledge measures the percentage of correct questions on the teacher content knowledge test (language for language teachers and mathematics for mathematics teachers), ranging from 0 to 100. Panel C includes both pedagogical skill and content knowledge, as described earlier, while panel D includes pedagogical skills sub-scores. In all columns, the model controls for a vector of student, teacher, and school characteristics. Student controls include age and gender. Teacher controls include gender and teaching experience in years. School controls include rurality, number of students enrolled, average teacher absence rate, infrastructure index, and a school-level wealth index. The infrastructure index measures whether a school has access to drinking water, functioning toilet, functioning internet, and functioning electricity in the classroom, and whether the school is accessible to students with disabilities. These models also control for survey time, the number of months from the start of the academic year to implementation of survey instrument. All models include country fixed effects. Standard errors are clustered by school and reported in parentheses. * $p< 0.1$, ** $p< 0.05$, *** $p< 0.01$

Table C5: Regression Results for Extended Model (By Income Group)

	(1) Student knowledge (Low income)	(2) Literacy student knowledge (Low income)	(3) Math student knowledge (Low income)	(4) Student knowledge (Lower-middle income)	(5) Literacy student knowledge (Lower-middle income)	(6) Math student knowledge (Lower-middle income)	(7) Student knowledge (Upper-middle income)	(8) Literacy student knowledge (Upper-middle income)	(9) Math student knowledge (Upper-middle income)
Panel A: Teacher pedagogical skill									
Teach score	1.03	0.0077	2.05	1.64**	2.49***	0.80	-0.15	2.04	-2.33
	(1.19)	(1.40)	(1.27)	(0.82)	(0.90)	(0.95)	(1.56)	(1.35)	(2.26)
Observations	18218	18218	18218	12466	12466	12466	3946	3946	3946
R^2	0.236	0.275	0.127	0.198	0.242	0.108	0.266	0.245	0.168
Panel B: Teacher content knowledge									
Content knowledge	0.065***	0.053**	0.076***	0.040	0.028	0.051	0.11***	0.032	0.18***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)
Observations	16292	16292	16292	12152	12152	12152	3790	3790	3790
R^2	0.245	0.295	0.123	0.192	0.238	0.106	0.281	0.231	0.199
Panel C: Pedagogical skill and content knowledge									
Teach score	0.54	-0.38	1.47	1.63*	2.34***	0.92	-0.43	1.91	-2.77
	(1.22)	(1.41)	(1.34)	(0.84)	(0.88)	(0.96)	(1.56)	(1.38)	(2.26)
Content knowledge	0.063***	0.055**	0.071**	0.038	0.025	0.050	0.11***	0.037	0.18***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)
Observations	16292	16292	16292	12152	12152	12152	3717	3717	3717
R^2	0.245	0.295	0.124	0.195	0.241	0.106	0.278	0.239	0.195
Panel D: Pedagogical skill sub-scores									
Classroom culture	0.24	-0.33	0.82	0.21	0.69	-0.27	-1.29	-0.33	-2.26
	(1.01)	(1.23)	(1.11)	(0.94)	(1.02)	(1.05)	(2.20)	(1.85)	(3.11)
Instruction	-0.94	-0.83	-1.05	1.65	1.27	2.02*	0.14	1.36	-1.08
	(1.03)	(1.30)	(1.22)	(1.00)	(1.08)	(1.14)	(1.62)	(1.39)	(2.47)
Socio emotional skills	1.98*	1.36	2.59**	-0.20	0.54	-0.93	0.71	0.63	0.78
	(1.10)	(1.39)	(1.19)	(0.95)	(1.03)	(1.05)	(2.00)	(1.72)	(3.00)
Observations	18218	18218	18218	12466	12466	12466	3946	3946	3946
R^2	0.237	0.276	0.129	0.199	0.242	0.110	0.267	0.246	0.169
School characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table presents point estimates from regressing student learning (student knowledge score) on dimensions of teaching quality using our extended model for low-income (Columns 1-3), lower-middle-income (Columns 4-6), and upper-middle-income countries (Columns 7-9). Columns 1, 4, and 7 refer to overall student knowledge, columns 2, 5, and 8 refer to student knowledge in literacy and columns 3, 6, and 9 refer to student knowledge in mathematics. In the extended model, we add variables as controls based on the outcome of the CIF, random forest and lasso models. In panel A, teacher pedagogical skill is measured by the TEACH score (1-5), an average of all items in the TEACH instrument. In panel B, teacher content knowledge measures the percentage of correct questions on the teacher content knowledge test (language for

language teachers and mathematics for mathematics teachers), ranging from 0 to 100. Panel C includes both pedagogical skill and content knowledge, as described earlier, while panel D includes pedagogical skills sub-scores. In all columns, the model controls for a vector of student, teacher and school characteristics. Student controls include age, gender, and student attendance. Teacher controls include gender, teaching experience in years, their highest level of education, and the time spent (in minutes) on reading practice. School controls include rurality, number of students enrolled, average teacher absence rate, infrastructure index, a school-level wealth index, operational management, first grade student knowledge and availability of textbooks. The infrastructure index measures if a school has access to drinking water, functioning toilet, internet, electricity in the classroom and how accessible the school is to students with disabilities. Operational management is a score of how well principals/head teachers responded to two vignettes on school management. These models also control for survey time, the number of months from the start of the academic year to implementation of survey instrument. All models include country fixed effects. Standard errors clustered by school in parentheses. * p<$ 0.1, ** p <$ 0.05, *** p <$ 0.01

Table C6: Regression Results for Extended Model (By Country)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Ethiopia	Niger	Rwanda	Sierra Leone	Chad	Jordan	Pakistan ICT	Pakistan KP	Pakistan Balochistan	Pakistan Sindh	Pakistan Punjab	Peru	Gabon
Panel A: Teacher pedagogical skill													
Teach score	0.17	-1.73	1.59	1.19	1.40	-0.40	4.50**	5.40***	2.61	1.50	2.41	1.49	0.0032
	(2.11)	(3.03)	(3.56)	(3.36)	(1.78)	(1.07)	(1.74)	(2.05)	(4.46)	(2.57)	(1.96)	(1.47)	(3.35)
Observations	4461	3524	3761	2813	3659	2504	972	2825	1159	1656	3350	2029	1917
R2	0.144	0.307	0.182	0.256	0.243	0.073	0.175	0.222	0.238	0.165	0.178	0.271	0.167
Panel B: Teacher content knowledge													
Content knowledge	0.0041	0.027	0.074	0.038	0.12***	-0.011	-0.063	0.054	-0.013	-0.092	0.24***	0.11***	0.14**
	(0.04)	(0.05)	(0.07)	(0.05)	(0.05)	(0.05)	(0.07)	(0.05)	(0.10)	(0.08)	(0.05)	(0.04)	(0.06)
Observations	4359	3453	2096	2776	3608	2385	972	2800	1159	1486	3350	1917	1873
R2	0.144	0.309	0.217	0.256	0.259	0.072	0.169	0.201	0.236	0.168	0.213	0.286	0.188
Panel C: Pedagogical skill and content knowledge													
Teach score	0.60	-1.74	1.63	1.25	1.04	-0.54	6.54***	5.25**	2.81	0.69	3.07*	0.99	0.45
	(2.24)	(3.06)	(4.61)	(3.40)	(1.73)	(1.09)	(2.41)	(2.06)	(4.47)	(2.70)	(1.83)	(1.53)	(3.26)
Content knowledge	0.000063	0.030	0.066	0.038	0.12**	-0.0095	-0.14*	0.043	-0.025	-0.089	0.25***	0.10***	0.14**
	(0.05)	(0.05)	(0.07)	(0.05)	(0.05)	(0.05)	(0.08)	(0.05)	(0.11)	(0.09)	(0.05)	(0.04)	(0.06)
Observations	4359	3453	2096	2776	3608	2385	972	2800	1159	1486	3350	1917	1800
R2	0.144	0.309	0.217	0.256	0.260	0.072	0.184	0.216	0.239	0.168	0.216	0.287	0.190
Panel D: Pedagogical skill sub-scores													
Classroom culture	-2.65	-1.85	3.02	-2.31	1.76	1.81	-0.88	-0.49	1.77	-3.90	3.15	-0.83	-2.56
	(2.16)	(2.12)	(2.10)	(1.77)	(1.91)	(1.67)	(1.41)	(2.07)	(3.88)	(3.18)	(2.06)	(1.80)	(4.78)
Instruction	0.63	-3.49	-1.71	3.12	-2.01	-2.81	5.22**	5.27**	5.34*	5.89*	-2.47	3.16**	-1.14
	(2.27)	(2.39)	(2.07)	(2.94)	(1.87)	(1.73)	(2.36)	(2.43)	(3.07)	(3.09)	(2.08)	(1.49)	(3.09)
Socio emotional skills	2.09	5.14*	0.35	0.90	1.49	0.51	-0.84	-0.26	-5.93	-1.42	1.97	-2.22	3.68
	(2.40)	(2.90)	(2.48)	(2.81)	(2.13)	(1.50)	(1.96)	(2.10)	(4.26)	(2.38)	(2.41)	(1.84)	(4.06)
Observations	4461	3524	3761	2813	3659	2504	972	2825	1159	1656	3350	2029	1917
R2	0.148	0.316	0.187	0.261	0.247	0.079	0.185	0.230	0.258	0.180	0.184	0.277	0.171
School characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Teacher characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Extended controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table presents point estimates from regressing student learning (student knowledge score) on dimensions of teaching quality using our extended model for all sampled education systems. All columns report the effect on overall student knowledge. In the extended model, we add variables as controls based on the outcome of the CIF, random forest and lasso models. In panel A, teacher pedagogical skill is measured by the TEACH score (1-5), an average of all items in the TEACH instrument. In panel B, teacher content knowledge measures the percentage of correct questions on the teacher content knowledge test (language for language teachers and mathematics for mathematics teachers), ranging from 0 to 100. Panel C includes both pedagogical skill and content knowledge, as described earlier, while panel D includes pedagogical skills sub-scores. In all columns, the model controls for a vector of student, teacher and school characteristics. Student controls include age, gender, and student attendance. Teacher controls include gender, teaching experience in years, their highest level of education, and the time spent (in minutes) on reading practice. School controls include rurality, number of students enrolled, average teacher absence rate, infrastructure index, a school-level wealth index, operational management, first grade student knowledge and availability of textbooks. The infrastructure index measures if a school has access to drinking water, functioning toilet, internet, electricity in the classroom and how accessible the school is to students with disabilities. Operational management is a score of how well principals/head teachers responded to two vignettes on school management. These models also control for survey time, the number of months from the start of the academic year to implementation of survey instrument. All models include country fixed effects. Standard errors clustered by school in parentheses. * p<$ 0.1, ** p <$ 0.05, *** p <$ 0.01

Table C7: Regression Results for Core PLAY Model (By Country)

	(1) Ethiopia (Overall)	(2) Ethiopia (Literacy)	(3) Ethiopia (Math)	(4) Sierra Leone (Overall)	(5) Sierra Leone (Literacy)	(6) Sierra Leone (Math)	(7) Peru (Overall)	(8) Peru (Literacy)	(9) Peru (Math)
Panel A: Teacher PLAY score									
PLAY score	21.2***	32.0***	10.4*	8.87***	12.7***	5.00	5.45	3.52	7.38
	(5.56)	(6.59)	(5.88)	(3.38)	(4.03)	(3.70)	(4.23)	(3.22)	(6.34)
Observations	4301	4301	4301	3042	3042	3042	1940	1940	1940
R^2	0.098	0.106	0.069	0.218	0.223	0.118	0.173	0.131	0.144
Panel B: Teacher pedagogical skill									
Teach score	0.38	-0.42	1.17	2.29	2.99	1.59	1.89	3.72***	0.068
	(2.31)	(2.68)	(2.34)	(3.29)	(3.88)	(3.55)	(1.28)	(0.99)	(1.93)
Observations	4301	4301	4301	3042	3042	3042	1940	1940	1940
R^2	0.079	0.074	0.066	0.204	0.204	0.114	0.173	0.143	0.140
Panel C: Pedagogical skill and content knowledge									
PLAY score	23.3***	35.6***	11.1*	8.96**	13.3***	4.59	-0.25	-3.83	3.32
	(5.95)	(7.18)	(6.11)	(3.45)	(4.06)	(3.82)	(4.74)	(3.58)	(7.37)
Teach score	-1.81	-4.13	0.52	0.55	0.31	0.80	1.23	3.80***	-1.35
	(2.56)	(2.85)	(2.65)	(3.41)	(3.98)	(3.72)	(1.40)	(1.21)	(2.08)
Content knowledge	0.025	0.061	-0.011	-0.0090	0.022	-0.040	0.085**	0.058**	0.11
	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)	(0.04)	(0.03)	(0.07)
Observations	4199	4199	4199	2980	2980	2980	1828	1828	1828
R^2	0.103	0.115	0.070	0.217	0.226	0.117	0.186	0.145	0.158
Panel D: PLAY skill sub-scores									
Support for exploration	8.43**	11.4***	5.42	4.63	6.30	2.96	1.96	1.66	2.26

	(3.29)	(3.70)	(3.60)	(3.52)	(4.58)	(3.60)	(2.01)	(1.62)	(2.94)
Support for exploration	5.68	3.50	7.86	-2.59	-9.78*	4.61	6.45	3.32	9.57
	(6.99)	(8.55)	(6.84)	(4.53)	(5.17)	(5.09)	(4.15)	(3.19)	(6.26)
Support for connection	8.22	1.85	14.6*	6.86*	12.0***	1.71	-4.07	-3.05	-5.10
	(7.75)	(8.85)	(7.85)	(3.85)	(4.13)	(4.15)	(2.72)	(2.08)	(4.24)
Support for emotional climate	2.12	7.61*	-3.38	-3.32	0.17	-6.82	2.24	1.81	2.68
	(3.60)	(4.23)	(3.72)	(4.62)	(5.40)	(5.00)	(3.01)	(2.75)	(4.37)
Observations	4301	4301	4301	3042	3042	3042	1940	1940	1940
R^2	0.100	0.108	0.077	0.229	0.242	0.128	0.181	0.135	0.151
Panel E: Teach skill sub-scores									
Classroom culture	-4.06*	-4.42	-3.70*	-1.59	-3.47	0.28	1.73	2.15	1.30
	(2.33)	(2.89)	(2.09)	(2.15)	(2.56)	(2.50)	(2.04)	(1.84)	(2.81)
Instruction	1.10	-1.08	3.29	3.72	6.47	0.98	0.15	2.75**	-2.46
	(2.25)	(2.52)	(2.31)	(3.26)	(3.93)	(3.68)	(1.44)	(1.16)	(2.10)
Socio emotional skills	3.10	5.37*	0.83	0.58	0.82	0.34	0.54	-1.42	2.51
	(2.31)	(2.76)	(2.37)	(3.07)	(3.52)	(3.47)	(1.95)	(1.38)	(3.03)
Observations	4301	4301	4301	3042	3042	3042	1940	1940	1940
R^2	0.089	0.087	0.073	0.208	0.211	0.114	0.174	0.147	0.143
School characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Student characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Extended controls	No	No	No	No	No	No	No	No	No
Survey time	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table presents point estimates from regressing student learning (student knowledge score) on dimensions of teaching quality using our core model for all PLAY countries. In panel A, teacher pedagogical skill is measured by the Play score (0-2), an average of all items in the Play instrument. In panel B, teacher pedagogical skill is measured by the TEACH score (1-5), an average of all items in the TEACH instrument. In panel C, teacher content knowledge is added which measures the percentage of correct questions on the teacher content knowledge test (language for language teachers and mathematics for mathematics teachers), ranging from 0 to 100. In all columns, the model controls for a vector of student, teacher, and school characteristics. Student controls include age and gender. Teacher controls include gender, teaching experience in years, and teacher's level of education. School controls include rurality, number of students enrolled, average teacher absence rate, infrastructure index, and a school-level wealth index. The infrastructure index measures whether a school has access to drinking water, functioning toilet, functioning internet, and functioning electricity in the classroom, and whether the school is accessible to students with disabilities. These models also control for survey time, the number of months from the start of the academic year to implementation of survey instrument. All models include country fixed effects. Standard errors clustered by school in parentheses. * p< 0.1, ** p< 0.05, *** p< 0.01

Table C8: Regression Results Teacher Pedagogical Skill (By Income Group)

	(1)	(2)	(3)
	Teach Score (Low income)	Teach score (Lower-middle income)	Teach score (Upper-middle income)
Female teacher	0.00088	0.098**	0.10
	(0.03)	(0.05)	(0.06)
Teaching experience (in years)	-0.0025	-0.0014	0.0038
	(0.00)	(0.00)	(0.00)
Highest education: Bachelors	0.073*	0.12	-0.084
	(0.04)	(0.09)	(0.14)
Highest education: master or doctoral	-0.069	0.15*	-0.030
	(0.14)	(0.09)	(0.20)
Highest education: Other	0.036	-0.018	-0.13
	(0.04)	(0.19)	(0.16)
Rurality	-0.013	-0.13**	0.025
	(0.03)	(0.05)	(0.07)
Infrastructure index	0.0041	0.063**	0.010
	(0.01)	(0.03)	(0.03)
Students enrolled	0.000063**	0.00012**	0.00040**
	(0.00)	(0.00)	(0.00)
Average teacher absence rate (school level)	-0.000053	0.00023	-0.0029*
	(0.00)	(0.00)	(0.00)
Recruitment score (0-9)	0.015*	-0.050***	-0.0033
	(0.01)	(0.02)	(0.01)
Probationary period for new teachers	0.012	-0.052	-0.16**
	(0.03)	(0.07)	(0.07)
Participated in induction/mentorship program	0.043	0.024	0.041
	(0.03)	(0.06)	(0.07)
Teaching practicum under pre-service training	-0.021	0.0021	-0.025
	(0.05)	(0.10)	(0.10)
Hours per day spent teaching during teaching practicum	0.00079	0.010	0.00073
	(0.01)	(0.03)	(0.02)
Participated in in-service training	0.038	-0.0088	0.016
	(0.04)	(0.06)	(0.08)
Days in-service training lasted	-0.0015	-0.00018	-0.000068
	(0.00)	(0.00)	(0.00)
Less than 50% in service training in classroom	-0.035	-0.079	0.095
	(0.05)	(0.14)	(0.11)
Over 50% in service training in classroom	0.075	0.064	0.095
	(0.07)	(0.15)	(0.11)
Classroom has been observed	-0.051	0.074	-0.0048
	(0.05)	(0.07)	(0.10)
Months since last classroom observation	0.0051	-0.0037	0.0022
	(0.00)	(0.01)	(0.01)
Discussed results of observation	0.12*	0.0037	-0.088
	(0.07)	(0.12)	(0.20)
Observer provided feedback	-0.10	-0.043	0.057
	(0.07)	(0.11)	(0.19)
Had lesson plans for last week	0.0072	-0.075	0.043
	(0.04)	(0.06)	(0.13)
Discussed lesson plans	0.045	0.034	0.13**
	(0.04)	(0.06)	(0.06)
Evaluated on pedagogical skill and content knowledge	-0.032	-0.027	-0.13*
	(0.03)	(0.05)	(0.07)
Financial consequences to	0.035	-0.0066	-0.13

negative evaluation			
	(0.04)	(0.06)	(0.11)
Financial consequences to positive evaluation	-0.024	-0.15**	-0.014
	(0.03)	(0.07)	(0.07)
Meritocracy for advancement	-0.022	0.017	0.068
	(0.03)	(0.06)	(0.07)
Received bonus during last AY	0.097**	0.050	0.075
	(0.05)	(0.14)	(0.14)
Salary was delayed during last AY	-0.056*	-0.12	-0.013
	(0.03)	(0.10)	(0.10)
Motivation - like teaching	-0.24*	-0.12**	0.019
	(0.13)	(0.06)	(0.07)
Motivation - steady career path	-0.21	-0.12	0.13
	(0.13)	(0.09)	(0.14)
Motivation - teaching benefits others	-0.16		
	(0.13)		
Intrinsic motivation (school level)	-0.0047	0.12	0.11
	(0.04)	(0.08)	(0.09)
Teacher monitoring (school level)	-0.043*	-0.051	0.019
	(0.03)	(0.04)	(0.05)
Teaching evaluation (school level)	0.0074	0.039	0.086*
	(0.02)	(0.03)	(0.05)
Teacher support (school level)	0.017	0.071*	-0.050
	(0.03)	(0.04)	(0.07)
Teacher attraction (school level)	0.013	0.12	-0.033
	(0.04)	(0.08)	(0.09)
Instructional leadership (school level)	0.0092	0.081**	-0.0029
	(0.02)	(0.04)	(0.05)
Constant	2.60***	1.59***	2.31***
	(0.25)	(0.35)	(0.49)
Observations	1017	810	265
R2	0.115	0.274	0.328
School characteristics	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes
Recruitment and Training	Yes	Yes	Yes
Instructional leadership	Yes	Yes	Yes
Evaluation	Yes	Yes	Yes
Incentives	Yes	Yes	Yes
Intrinsic motivation	Yes	Yes	Yes
School level indicators	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes

For the highest level of education completed, the omitted category is high school and below. For the percent of in-service training done in classroom, the omitted category is that no in-service training was done in the classroom. A recruitment score sums up the number of questions a teacher responded yes to under the recruitment category (table 8). An infrastructure index measures if a school has access to drinking water, functioning toilet, internet, electricity in the classroom and how accessible the school is to students with disabilities.* p<0.01 ** p<0.05 *** p<0.10. Standard errors clustered at school level in parentheses.

Table C9: Regression Results – Teacher Content Knowledge (By Income Group)

	(1)	(2)	(3)
	Content knowledge (Low income)	Content knowledge (Lower-middle income)	Content knowledge (Upper-middle income)
Female teacher	-3.00***	1.57*	-2.17
	(0.72)	(0.92)	(1.38)
Teaching experience (in years)	0.051	0.018	-0.20**
	(0.04)	(0.04)	(0.08)

Highest education: Bachelors	3.37***	2.52*	10.1***
	(1.27)	(1.52)	(2.62)
Highest education: master or doctoral	0.62	2.56*	12.5***
	(3.07)	(1.46)	(3.24)
Highest education: Other	3.28***	0.88	5.42*
	(1.11)	(2.65)	(3.02)
Rurality	-2.91***	-1.41	0.95
	(1.10)	(0.95)	(2.19)
Infrastructure index	1.84***	-0.37	1.28
	(0.45)	(0.50)	(0.97)
Students enrolled	-0.0037***	-0.00088	0.0060*
	(0.00)	(0.00)	(0.00)
Average teacher absence rate (school level)	-0.030	0.051	0.056
	(0.02)	(0.03)	(0.05)
Recruitment score (0-9)	-0.29	-0.23	1.36***
	(0.24)	(0.29)	(0.40)
Probationary period for new teachers	-0.067	-0.43	-2.64
	(0.83)	(0.98)	(1.63)
Participated in induction/mentorship program	-0.36	0.041	-4.49***
	(0.76)	(0.81)	(1.66)
Teaching practicum under pre-service training	1.92	-0.30	0.12
	(1.30)	(1.52)	(2.29)
Hours per day spent teaching during teaching practicum	-0.26	-0.51	-0.14
	(0.24)	(0.37)	(0.41)
Participated in in-service training	-2.57**	-0.97	0.14
	(1.00)	(1.03)	(1.81)
Days in-service training lasted	0.10*	0.043	0.17***
	(0.06)	(0.05)	(0.06)
Less than 50% in service training in classroom	2.36*	1.97	-3.13
	(1.32)	(1.79)	(2.57)
Over 50% in service training in classroom	6.19***	1.38	-1.33
	(1.72)	(1.92)	(2.37)
Classroom has been observed	-1.85	-1.17	0.60
	(1.17)	(1.06)	(2.91)
Months since last classroom observation	-0.0032	-0.029	-0.18
	(0.00)	(0.17)	(0.13)
Discussed results of observation	1.98	0.78	-3.50
	(1.57)	(1.52)	(4.62)
Observer provided feedback	-1.80	-0.042	4.71
	(1.39)	(1.41)	(4.10)
Had lesson plans for last week	2.38**	2.31**	-3.17
	(1.04)	(1.01)	(2.78)
Discussed lesson plans	-0.64	-0.098	1.42
	(0.87)	(0.89)	(1.57)
Evaluated on pedagogical skill and content knowledge	-1.90**	0.95	-5.50***
	(0.80)	(0.88)	(1.67)
Financial consequences to negative evaluation	1.73	0.61	0.50
	(1.19)	(0.97)	(2.48)
Financial consequences to positive evaluation	1.52*	-1.59*	1.62
	(0.82)	(0.93)	(1.69)
Meritocracy for advancement	-0.40	0.071	0.089
	(0.85)	(0.88)	(1.68)
Received bonus during last AY	1.74	-3.48**	4.73*
	(1.11)	(1.58)	(2.54)
Salary was delayed during	-1.60*	3.23**	-1.97

last AY			
	(0.86)	(1.36)	(2.32)
Motivation - like teaching	5.00**	1.71	1.10
	(2.31)	(1.10)	(1.43)
Motivation - steady career path	6.11**	3.66***	2.73
	(2.48)	(1.33)	(3.55)
Motivation - teaching benefits others	6.79***		
	(2.36)		
Intrinsic motivation (school level)	2.06*	-1.48	2.97
	(1.19)	(1.22)	(2.83)
Teacher monitoring (school level)	-0.33	1.37*	-2.19
	(0.83)	(0.79)	(1.66)
Teaching evaluation (school level)	0.65	0.29	2.17
	(0.82)	(0.66)	(1.61)
Teacher support (school level)	-0.32	0.46	1.61
	(0.79)	(0.72)	(1.67)
Teacher attraction (school level)	-0.75	3.37**	-5.63**
	(1.24)	(1.31)	(2.78)
Instructional leadership (school level)	0.43	-0.66	2.45**
	(0.62)	(0.68)	(1.25)
Constant	40.3***	46.8***	46.4***
	(7.00)	(6.19)	(14.77)
Observations	4483	3782	1151
R2	0.121	0.256	0.146
School characteristics	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes
Recruitment and Training	Yes	Yes	Yes
Instructional leadership	Yes	Yes	Yes
Evaluation	Yes	Yes	Yes
Incentives	Yes	Yes	Yes
Intrinsic motivation	Yes	Yes	Yes
School level indicators	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes

For the highest level of education completed, the omitted category is high school and below. For the percent of in-service training done in classroom, the omitted category is that no in-service training was done in the classroom. A recruitment score sums up the number of questions a teacher responded yes to under the recruitment category (table 8). An infrastructure index measures if a school has access to drinking water, functioning toilet, internet, electricity in the classroom and how accessible the school is to students with disabilities.* p<0.01 ** p<0.05 *** p<0.10. Standard errors clustered at school level in parentheses.

Table C10: Regression Results – Teacher Pedagogical skills for the Sample of PLAY Countries Only

	(1) Teach score	(2) Teach score	(3) Teach score	(4) Teach score	(5) Teach score	(6) Teach score	(7) Teach score
Female teacher	0.071	0.071	0.063	0.066	0.064	0.065	0.074
	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)
Teaching experience (in years)	0.00048	0.00032	-0.00032	-0.00037	-0.00041	-0.00053	-0.00010
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Highest education: Bachelors	0.16**	0.17**	0.16**	0.16**	0.17**	0.17**	0.15**
	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Highest education: master or doctoral	0.33***	0.30**	0.29**	0.30**	0.30**	0.26**	0.24*
	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)	(0.13)	(0.13)
Highest education: Other	0.017	0.023	0.030	0.031	0.037	0.025	-0.0015
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Rurality	-0.075*	-0.072*	-0.076*	-0.079*	-0.074*	-0.064	-0.062
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Infrastructure index	-0.0098	-0.010	-0.017	-0.017	-0.017	-0.018	-0.017
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Students enrolled	0.000054	0.000054	0.000051	0.000051	0.000053	0.000066	0.000060
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Average teacher absence rate	-0.0024**	-0.0025**	-0.0023**	-0.0022**	-0.0021**	-0.0022**	-0.0021**

(school level)							
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Recruitment score (0-9)		-0.0081	-0.012	-0.011	-0.0096	-0.0075	-0.011
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Probationary period for new teachers		-0.036	-0.032	-0.030	-0.024	-0.023	-0.023
		(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
Participated in induction/mentorship program		0.023	0.027	0.029	0.030	0.026	0.051
		(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
Teaching practicum under pre-service training		-0.030	-0.035	-0.038	-0.042	-0.054	-0.034
		(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Hours per day spent teaching during teaching practicum		0.0073	0.0056	0.0058	0.0064	0.0077	0.0073
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Participated in in-service training		0.044	0.044	0.047	0.050	0.050	0.062
		(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Days in-service training lasted		-0.00039	-0.00054	-0.00051	-0.00036	-0.00040	-0.00043
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Less than 50% in service training in classroom		0.044	0.040	0.041	0.039	0.035	0.036
		(0.07)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)
Over 50% in service training in classroom		0.063	0.071	0.071	0.076	0.085	0.085
		(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Classroom has been observed			-0.054	-0.052	-0.053	-0.060	-0.077
			(0.06)	(0.07)	(0.07)	(0.07)	(0.07)
Months since last classroom observation			0.011**	0.011**	0.011**	0.011**	0.011**
			(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Discussed results of observation			0.20**	0.20*	0.18*	0.18*	0.13
			(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
Observer provided feedback			-0.20**	-0.20**	-0.19**	-0.17**	-0.16*
			(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Had lesson plans for last week			-0.055	-0.052	-0.046	-0.037	-0.029
			(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Discussed lesson plans			0.14***	0.14***	0.14***	0.14***	0.12**
			(0.04)	(0.04)	(0.04)	(0.04)	(0.05)
Evaluated on pedagogical skill and content knowledge				-0.0064	-0.012	-0.0084	-0.040
				(0.04)	(0.04)	(0.04)	(0.04)
Financial consequences to negative evaluation				0.021	0.016	0.025	0.026
				(0.06)	(0.06)	(0.06)	(0.06)
Financial consequences to positive evaluation				-0.032	-0.025	-0.025	-0.032
				(0.04)	(0.04)	(0.04)	(0.04)
Meritocracy for advancement					-0.034	-0.044	-0.040
					(0.05)	(0.05)	(0.05)
Received bonus during last AY					0.043	0.055	0.054
					(0.10)	(0.10)	(0.10)
Salary was delayed during last AY					-0.045	-0.045	-0.042
					(0.05)	(0.05)	(0.05)
Motivation - like teaching						-0.23**	-0.25**
						(0.11)	(0.11)
Motivation - steady career path						-0.19	-0.20
						(0.13)	(0.13)
Motivation - teaching benefits others						-0.14	-0.16
						(0.12)	(0.12)
Intrinsic motivation (school level)							0.014
							(0.06)
Teacher monitoring (school							-0.019

level)							
							(0.03)
Teaching evaluation (school level)							0.049
							(0.03)
Teacher support (school level)							-0.041
							(0.03)
Teacher attraction (school level)							0.024
							(0.06)
Instructional leadership (school level)							0.033
							(0.03)
Constant	2.43***	2.44***	2.44***	2.45***	2.46***	2.64***	2.42***
	(0.08)	(0.08)	(0.10)	(0.10)	(0.11)	(0.15)	(0.32)
Observations	566	566	566	566	566	566	565
R^2	0.217	0.224	0.247	0.248	0.251	0.259	0.267
School characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Recruitment and Training	No	Yes	Yes	Yes	Yes	Yes	Yes
Instructional leadership	No	No	Yes	Yes	Yes	Yes	Yes
Evaluation	No	No	No	Yes	Yes	Yes	Yes
Incentives	No	No	No	No	Yes	Yes	Yes
Intrinsic motivation	No	No	No	No	No	Yes	Yes
School level indicators	No	No	No	No	No	No	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

For the highest level of education completed, the omitted category is high school and below. For the percent of in-service training done in classroom, the omitted category is that no in-service training was done in the classroom. A recruitment score sums up the number of questions a teacher responded yes to under the recruitment category (table 8). An infrastructure index measures if a school has access to drinking water, functioning toilet, internet, electricity in the classroom and how accessible the school is to students with disabilities.* p<0.01 ** p<0.05 *** p<0.10. Standard errors clustered at school level in parentheses.

ABSTRACT

Using unique nationally representative school and system survey data from 13 education systems in low- and middle-income countries collected through the World Bank's Global Education Policy Dashboard (GEPD), we examine how the pedagogical practices, including practices to foster student engagement, and subject content knowledge of primary-school teachers correlate with their students' learning outcomes. We find that student performance on literacy (and to a lesser extent math) assessments is correlated with receiving instruction from teachers with better measured pedagogical skills. While the better-pedagogy effect is modest for the full sample, it is statistically robust and quite substantial for the upper-middle-income countries in the sample. Based on a sub-sample of those education systems, we also find that the use of learning strategies that support greater student engagement appears to be highly predictive of student learning outcomes in literacy. Better pedagogical practices are correlated with teachers' exposure to more practical, school-based pedagogical support—for example through induction or mentoring and feedback on lesson plans—and also with better teacher evaluation at the school level. The findings confirm the important role of interventions providing direct pedagogical support and feedback to teachers through training, instructional leadership, and evaluation, and they highlight the potential for interventions to foster student engagement and improve learning outcomes.