SERVICE DELIVERY INDICATORS

Education | Health

Education Service Delivery in TANZANIA

WORLD BANK GROUP

AFRICAN ECONOMIC RESEARCH CONSORTIUM
Consortium pour la Recherche Economique en Afrique

Tanzania 2014 Service Delivery Indicators

Education Technical Report

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

This report has been prepared in consultation with the Government of Tanzania by Waly Wane and Gayle Martin (Task Team Leaders). The World Bank team included Jud Shearer, together with the SDI team members Eleni Papakosta, Raihona Atakhodjayeva, and Christophe Rockmore. Data collection was undertaken by the Research on Poverty Alleviation (REPOA) under the leadership of Prof. Samwel Wangwe and Lucas Katera.

The team would like to use the opportunity to thank the officials of the Ministry of Education as well as the respondents who participated in the survey without whom this study would not have been possible. We would also like to thank the World Bank's Tanzania Country Office and especially Philip Dongier (Country Director), Rekha Menon (Program Leader), Arun Joshi (Education Task Team Leader), and Nkahiga Kaboko (Education Specialist) for their guidance and support.

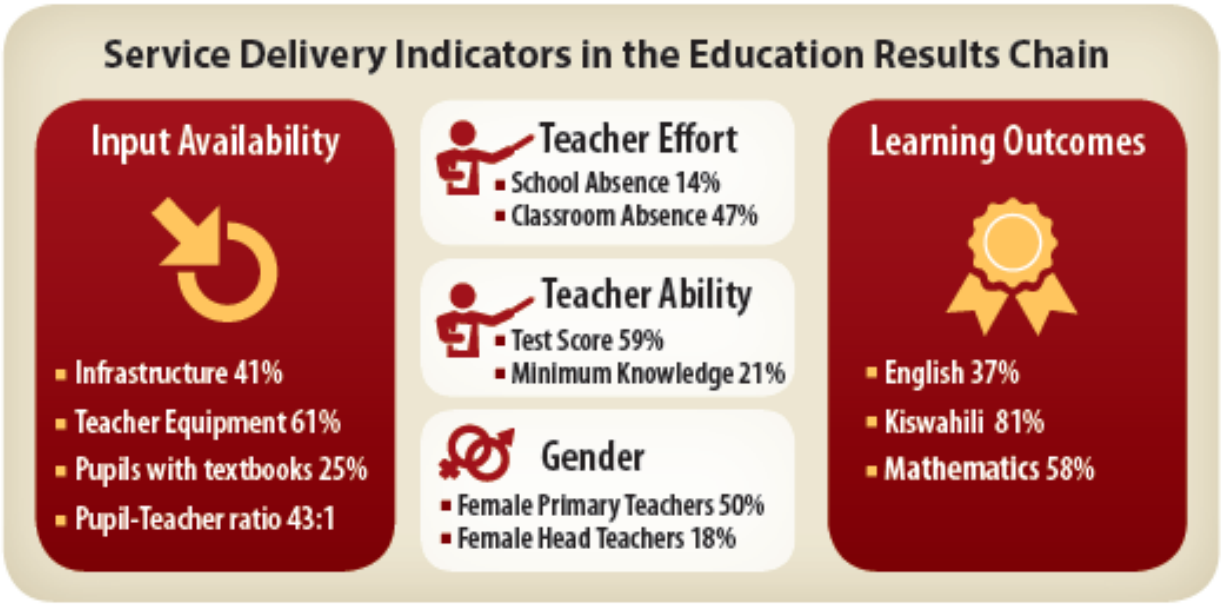
Finally, the team gratefully acknowledges the financial support from the DFID, William and Flora Hewlett Foundation, in addition to the resources from the World Bank.

EXECUTIVE SUMMARY

The Service Delivery Indicators provide a set of metrics for benchmarking service delivery performance in education and health. The overall objective of the indicators is to gauge the quality of service delivery in primary education and basic health services. The indicators enable the identification of gaps and tracking of progress over time and across countries. It is envisaged that the broad availability, high public awareness and a persistent focus on the indicators will mobilize policymakers, citizens, service providers, donors and other stakeholders for action to improve the quality of services and ultimately to improve development outcomes and social welfare.

This report presents the findings from the implementation of the Service Delivery Indicators in the Education sector in Tanzania in 2014. Survey implementation was preceded by an extensive consultation with Government and key stakeholders on survey design, sampling, and adaptation of survey instruments. Pre-testing of the survey instruments, enumerator training, and fieldwork took place in 2014.

Information was collected from 400 primary schools, 2,196 teachers (for skills assessment), 3,692 teachers (for absence rate) and 4,041 pupils across Tanzania. The results provide a snapshot of the quality of service delivery and the physical environment within which services are delivered in public primary schools. The survey provides information on (i) teacher effort; (ii) teacher knowledge and ability; and (iii) the availability of key inputs, such as textbooks, basic teaching equipment and infrastructure (such as sanitation or quality of lighting).



What providers do (teachers' effort)

On average, 14 percent of teachers were found to be absent from school. Absence from classroom was much more serious with almost half (46.5 percent) of the teachers not in the class teaching. While in the classroom, teachers spent on average about 12 percent of time on non-teaching activities.

Combining the absence from school and the classroom with the time engaged in non-teaching activities, the results indicate that pupils only had 2 hours and 47 minutes of teaching time every day out of the allocated 5 hours 55 minutes. It is worth noting that over 80 percent of teachers were absent with management approval, suggesting: (i) management weakness and a sub-optimal allocation of paid staff time; (ii) absence is within the power of management to influence, and (iii) absence may be amenable to action in the short run.

What providers know (teachers' knowledge and ability)

The vast majority of teachers lacked the necessary academic and pedagogical skills to teach. The average score on the mathematics and English assessment among teachers who teach either or both subjects was 49.5 percent. Only 21.4 percent of the teachers scored at least 80 percent on these assessments. Pupils cannot learn more from their teachers than what the teachers know, and, therefore, teachers' lack of technical competences severely constrained learning outcomes in Tanzania.

What providers have (availability of key inputs)

The pupil-teacher ratio averaged 43.4 pupils per teacher, slightly below the expected norm of 45:1. However, significant gaps existed in the availability of inputs at the frontline. Only 40.9 percent of schools had the minimum infrastructure. Most striking was the absence of functional, improved, accessible, private, and clean toilets. Fewer than half (47 percent) of all primary schools surveyed had toilets (such as a ventilated improved pit latrine, etc.) meeting the standard. Only three out of five (61 percent) of the schools had the minimum teaching materials. Despite the recent nationwide distribution of textbooks, only 25 percent of the pupils had a mathematics or English textbook in the classroom.

Against this background, one must also note that teachers experienced salary delays with almost one out of three teachers (32.9 percent) claiming to have experienced such a delay at least once over the year preceding the survey. Teachers also complained about unpaid claims with half (46.7 percent) of them reporting at least one unpaid claim, this proportion reached 70.4 percent in Dar es Salaam. The most prevalent unpaid claim nationally was salary arrears (23.7 percent) followed by unpaid leave (9.4 percent) and relocation allowance and hardship pay with 3.5 percent and 3.4 percent respectively.

Tanzania SDI trends and how does Tanzania compare to other countries

Between 2010 and 2014, Tanzania made significant gains in terms of smaller class sizes, better teaching equipment, and higher teacher presence in school. As a result, the average Tanzanian primary pupil gained 43 minutes of teaching time per day while teaching time doubled for the urban pupils. On the other hand, teachers' subject and pedagogical content knowledge stagnated or slightly deteriorated.

Bearing in mind that overall performance is low, Tanzania fares relatively well among the SDI countries in terms of inputs, effort, and knowledge. Tanzania's schools have worse infrastructure compared to schools in its EAC neighbors, but they fare better than Nigeria, Togo, and Mozambique. Although only 1 in 4 (26 percent) of Tanzanian pupils used a textbook in the classroom, Tanzania outperformed Uganda where only 6 percent of the pupils used a textbook in the classroom. In contrast, 76 percent, 68 percent, 44 percent, and 34 percent of pupils in Togo, Mozambique, Kenya, and Nigeria, respectively used a textbook in class.

When it comes to teachers' effort, Tanzania and Kenya displayed patterns with relatively low school absence rates (15.3 percent and 15.2 percent, respectively), but high classroom absence rates with almost half of the teachers not found in classroom at any point in time. Mozambique and Uganda have significantly higher absence rates for both. The three West African countries (Nigeria, Senegal, and Togo) have much lower classroom absence rates.

Kenyan, Ugandan, and Tanzanian teachers significantly outperformed Nigerian, Togolese, and Mozambican teachers in minimum knowledge. Tanzania where 15.6 percent of teachers are assessed to have minimum knowledge to teach is second only to Kenya (34.8 percent), the best performer among SDI countries. Almost no teachers in Nigeria and Togo passed the 80 percent score bar to be considered as mastering the curriculum they teach. Tanzanian teachers' average test scores for English, mathematics, and pedagogy (46.6 percent) are also second only to their Kenyan counterparts (55.6 percent). Pupils' test scores mirror their teachers' performance and Tanzanian pupils are outperformed only by their Kenyan peers (49.2 percent and 69.4 percent, respectively).

Table 1. Comparison of SDI results across countries (public schools only)[1]

	Tanzania 2014	Average SDI	Kenya 2012	Mozambique 2014	Nigeria* 2013	Senegal 2011	Tanzania 2011	Togo 2013	Uganda 2013
Teacher Ability									
Minimum knowledge (At least 80% in language and mathematics)	21.5	12.7	34.8	0.3	2.4	Not Comparable	Not Comparable	0.9	19.4
Test score (language, mathematics, and pedagogy)	48.3	42.0	55.6	26.9	30.5	Not Comparable	Not Comparable	33.9	45.5
Teacher Effort									
School absence rate	14.4	20.1	15.2	44.8	16.9	18.0	23.0	22.6	29.9
Classroom absence rate	46.7	42.1	47.3	56.2	22.8	29.0	53.0	39.3	56.9
Scheduled teaching time	*5h 54min*	*5h 31min*	*5h 31min*	*4h 17min*	*4h 44min*	*4h 36min*	*5h 12min*	*5h 28min*	*7h 13min*
Time spent teaching per day	2h 46min	2h 53min	2h 30min	1h 41 min	3h 10min	3h 15min	2h 04min	3h 15min	2h 56min
Availability of Inputs									
Observed pupil-teacher ratio	43.5	42.1	39.3	21.4	21.5	27.2	52.0	31.4	53.9
Share of pupils with textbooks	25.3	37.2	44.5	68.1	33.7	18.0	19.7	76.0	6.0
Minimum equipment availability (90% with pencils and notebooks)	61.4	57.8	74.3	76.8	48.2	Not Comparable	Not Comparable	24.3	79.5
Minimum infrastructure availability	40.4	36.2	60.2	29.1	13.4	Not Comparable	Not Comparable	14.4	57.2
Pupil Learning									
Test Score (out of 100) (language, mathematics)	40.1+*	45.4	69.4	20.8	25.1	Not Comparable	Not Comparable	38.1	45.3
Language test score	36.5+*	44.8	72.5	18.7	23.3	Not Comparable	Not Comparable	36.9	43.4
Mathematics test score	58.2	45.2	57.4	25.1	28.2	Not Comparable	Not Comparable	41.3	41.7

Note: (*) Values for Nigeria are the weighted average of the four states surveyed, namely Anambra, Bauchi, Ekiti, and Niger.

[1] These numbers may be different from the previously published country reports because the methodology for calculating the indicators has been updated. The numbers shown here are current. To find out more about how the indictors are calculated, go to www.SDIndicators.org.

Table 2: Comparison of SDI results across countries (All schools)[2]

	Tanzania* 2014	Average SDI	Kenya 2012	Mozambique+ 2014	Nigeria** 2013	Senegal+ 2011	Tanzania+ 2011	Togo 2013	Uganda 2013
Teacher Ability									
Minimum knowledge (At least 80% in language and mathematics)	21.5	**14.6**	40.4	0.3	3.7	Not Comparable	Not Comparable	1.6	19.5
Test score (language, mathematics, and pedagogy)	48.3	**43.0**	57.1	26.9	32.9	Not Comparable	Not Comparable	35.6	45.3
Teacher Effort									
School absence rate	14.4	**18.6**	14.1	44.8	13.7	18.0	23.0	20.5	26.0
Classroom absence rate	46.7	**39.8**	42.1	56.2	19.1	29.0	53.0	35.8	52.8
Scheduled teaching time	*5h 54min*	***5h 34min***	*5h 37min*	*4h 17min*	*4h 53min*	*4h 36min*	*5h 12min*	*5h 29min*	*7h 18min*
Time spent teaching per day	2h 46min	**3h 02min**	2h 49min	1h 41 min	3h 26min	3h 15min	2h 04min	3h 29min	3h 18min
Availability of Inputs									
Observed pupil-teacher ratio	43.5	**40.4**	35.2	21.4	21.6	27.2	52.0	29.7	47.9
Share of pupils with textbooks	25.3	**37.1**	48.0	68.1	38.2	18.0	19.7	68.5	5.0
Minimum equipment availability (90% with pencils and notebooks)	61.4	**60.5**	78.8	76.8	54.8	Not Comparable	Not Comparable	26.4	80.6
Minimum infrastructure availability	40.4	**38.1**	59.5	29.1	18.5	Not Comparable	Not Comparable	22.3	53.7
Pupil Learning									
Test Score (out of 100) (language, mathematics)	40.1+*	**49.6**	72.0	20.8	32.2	Not Comparable	Not Comparable	45.7	48.6
Language test score	36.5+*	**49.5**	75.4	18.7	31.4	Not Comparable	Not Comparable	45.5	47.1
Mathematics test score	58.2	**47.3**	59.0	25.1	31.9	Not Comparable	Not Comparable	44.6	43.4

Note: (*)Because of the very low number of private schools, only public schools are included in this second round as well. (**)Values for Nigeria are the weighted average of the four states surveyed, namely Anambra, Bauchi, Ekiti, and Niger. (+) In Mozambique, Senegal, and Tanzania 2011 (round 1) only public schools were surveyed. (+*) These test scores are for pupils who were assessed in English., for those assessed in Kiswahili they scored 76.2 on the combined test and 80.9 on Kiswahili.

[2] A few of these numbers may be different from the previously published country reports because the methodology for calculating the indicators has been updated. The numbers shown here are current. To find out more about how the indictors are calculated, go to www.SDIndicators.org.

I. INTRODUCTION[3]

1. **Between May and September 2014, fourteen education teams travelled across Tanzania to collect data for the Service Delivery Indicators (SDI).** The SDI has been piloted in Senegal and Tanzania in 2010 and lessons learned from the pilot have led to a revised SDI being rolled out across Africa. Thus, this SDI follows a series of countries in Africa which have already implemented a full-fledged SDI (Kenya, Uganda, Nigeria, Togo, and Mozambique). Tanzania is, however, the first country to implement its second SDI which will allow for basic trend analysis. The Tanzania SDI has been implemented in both health and education sectors. This report concerns the education SDI.

2. **The Education SDI visited a representative sample of Tanzanian primary schools.** Following an initial consultation meeting drawing in several stakeholders, the initial SDI instruments have been customized to fit the Tanzanian context and it was decided that the SDI would also provide few sub-national indicators. Only public primary schools were including in this SDI.

3. The survey was implemented by Research in Poverty Alleviation (REPOA) in close coordination with a World Bank team. In each school, one standard four English or mathematics class was observed. Up to 10 pupils were randomly chosen amongst standard four learners and a total of 4,041 pupils were assessed for literacy and numeracy skills. For comparison with the 2010 SDI and other SDI countries, 2,841 pupils were tested in English. In addition, 1,200 more pupils were tested in Kiswahili, which is the major medium for learning in Tanzania's primary education system. Teachers also were assessed with 2,197 of them sitting through a 1 hour 10 minute assessment on their English, mathematics, and pedagogical skills. Finally, 3,692 teachers across grades were randomly chosen during the first visit and their whereabouts assessed in a second unannounced visit for estimation of teachers' effort and the level of absence in schools and classrooms.

4. The education service delivery indicators build on a growing body of literature on measuring the performance of schools and specifically teachers who are arguably the most important actors in the sector. The indicators provide a snapshot of the learning environment and key resources which need to be in place for pupils to learn. As expenditure on teachers represents by far the largest share of education spending in developing countries, and as several recent studies convincingly demonstrate how changes in teacher behavior can improve learning achievement, a strong focus is placed on the knowledge, skills and effort of teachers.

5. Annex B presents the Service Delivery Indicators in education and a short definition of each indicator. Below we give some more intuition for the choice of indicators, before presenting the results and a detailed discussion of the findings in the following sections.

6. A minimum requirement for learning is that the teachers are not absent from the school and spend time in the class rather than somewhere else. The first two indicators — *School absence rate* and *Classroom absence rate* — are direct measures of the extent to which this is the case. While having teachers in the class is a necessary condition, however, it is not sufficient for learning to take place. Teachers need to be involved in teaching and teachers need to have at least a minimum level of knowledge of the subjects they are teaching and skills to transform their knowledge into

[3] While not the focus of this report, the SDI surveys are also implemented in the health sector.

meaningful teaching. The indicators *share of teachers with minimum knowledge* measures to what extent these skills exist across schools and the indicator *time spent teaching in the class* the extent to which teachers are exerting effort to enable learning.

7. Four of the indicators deal with the environment i.e. the school. The indicator *availability of teaching resources* assesses if necessary materials such as blackboard, chalk, pencils, paper are in place to support the teaching activities. The indicator *school infrastructure* measures whether functioning sanitation facilities exist and if there is at least minimum light in the classroom so that pupils can read and study. Finally, the indicators *pupil-teacher ratio* and *share of pupils with textbooks* measure the average number of pupils per teacher in standard four classrooms and the proportion of pupils in the classroom who are using the relevant (mathematics or language) textbook while the teacher dispenses learning.

Table 3. 2014 Tanzania Service Delivery Indicators At-A-Glance

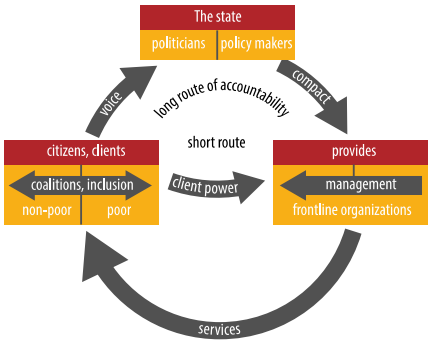
	Tanzania	Dar es Salaam	Other Urban	Rural	EQUIP-T	All Rural	All Urban
School absence rate (% of teachers)	14.4	17.8	13.7	14.3	8.3	14.3	14.6
Classroom absence rate (% of teachers)	46.7	42.9	46.7	47.0	43.5	47.0	45.8
Classroom teaching time (ToT)	2h 46min	2h 42min	2h 43min	2h 47min	2h 32min	2h 47min	2h 42min
Scheduled teaching time	*5h 56min*	*5h 41min*	*5h 50min*	*5h 57min*	*5h 58min*	*5h 57min*	*5h 48min*
Teachers' minimum knowledge	21.5	26.2	22.3	21.0	19.5	21.0	23.2
Observed pupil-teacher ratio	43.5	69.8	58.9	40.7	46.1	40.8	60.4
Share of pupils with textbooks	25.3	31.0	14.1	26.7	22.2	26.7	16.7
Minimum equipment availability	61.4	83.2	80.1	58.3	58.9	58.3	80.4
Minimum infrastructure availability	40.4	67.0	61.4	36.8	34.3	36.9	62.3
Infrastructure availability [a]	2.3	15.7	9.6	0.8	0.0	0.9	10.2

Note: (*a*) Comparable to SDI 2010 (i.e., school has electricity, toilet, and clean water). Source: Tanzania SDI 2014 and author's calculations.

Box 1. Analytical Underpinnings

Service delivery outcomes are determined by the relationships of accountability between policymakers, service providers, and citizens (Figure 1, World Bank 2004). Human development outcomes are the result of the interaction between various actors in the multi-step service delivery system, and depend on the characteristics and behavior of individuals and households. While delivery of quality education is contingent foremost on what happens in classrooms, a combination of several basic elements have to be present in order for quality services to be accessible and produced by teachers at the frontline, which depend on the overall service delivery system and supply chain. Adequate financing, infrastructure, human resources, material, and equipment need to be made available, while the institutions and governance structure provide incentives for the service providers to perform.

Figure 1. Relationships of accountability between citizens, service providers, and policymakers



Service Delivery Production Function

Consider a service delivery production function, f, which maps physical inputs, x, the effort put in by the service provider, e, as well as his/her type (or knowledge), θ, to deliver quality services into individual level outcomes, y. The effort variable e could be thought of as multidimensional and thus include effort (broadly defined) of other actors in the service delivery system. We can think of type as the characteristic (knowledge) of the individuals who are selected for a specific task. Of course, as noted above, outcomes of this production process are not just affected by the service delivery unit, but also by the actions and behaviors of households, which we denote by ε. We can therefore write

$$y = f(x,e,\theta) + \varepsilon$$

To assess the quality of services provided, one should ideally measure $f(x,e,\theta)$. Of course, it is notoriously difficult to measure all the arguments that enter the production, and would involve a huge data collection effort. A more feasible approach is therefore to focus instead on proxies of the arguments which, to a first-order approximation, have the largest effects.

Indicator Categories and the Selection Criteria

There are a host of data sets available in education. To a large extent, these data sets measure inputs and outcomes/outputs in the service delivery process, mostly from a household perspective. While providing a wealth of information, existing data sources (like Living Standards Measurement Survey (LSMS), Welfare Monitoring Surveys (WMS), and Core Welfare Indicators Questionnaire Survey (CWIQ)) cover only a sub-sample of countries and are, in many cases, outdated.

Box 1. Analytical Underpinnings (cont'd)

The proposed choice of indicators takes its starting point from the recent literature on the economics of education and service delivery, more generally. Overall, this literature stresses the importance of provider behavior and competence in the delivery of education services (as opposed to water and sanitation services and housing that rely on very different service delivery models). Conditional on service providers exerting effort, there is also some evidence that the provision of physical resources and infrastructure has important effects on the quality of service delivery.

The somewhat weak relationship between resources and outcomes documented in the literature has been associated with deficiencies in the incentive structure of school and education systems. Indeed, most service delivery systems in developing countries present frontline providers with a set of incentives that negate the impact of pure resource-based policies. Therefore, while resources alone appear to have a limited impact on the quality of education in developing countries, it is possible inputs are complementary to changes in incentives, so coupling improvements in both may have large and significant impacts (see Hanushek, 2006). As noted by Duflo, Dupas, and Kremer (2011), the fact that budgets have not kept pace with enrollment, leading to large pupil-teacher ratios, overstretched physical infrastructure, and insufficient number of textbooks, etc., is problematic. However, simply increasing the level of resources might not address the quality deficit in education without also taking providers' incentives into account.

SDI proposes three sets of indicators: (i) provider effort; (ii) knowledge of service providers and (iii) availability of key infrastructure and inputs at the frontline service provider level. Providing countries with detailed and comparable data on these important dimensions of service delivery is one of the main innovations of the Service Delivery Indicators.

Additional considerations in the selection of indicators are (i) quantitative (to avoid problems of perception biases that limit both cross-country and longitudinal comparisons), (ii) ordinal in nature (to allow within and cross-country comparisons); (iii) robust (in the sense that the methodology used to construct the indicators can be verified and replicated); (iv) actionable; and (v) cost effective to collect.

Table 4. Education Indicators

Teacher Effort
School absence rate
Classroom absence rate
Time spent teaching per day
Teacher Knowledge and Ability
Minimum knowledge in mathematics
Minimum knowledge in English
Minimum knowledge in pedagogy
Availability of Inputs
Minimum infrastructure availability
Minimum equipment availability
Share of pupils with textbooks
Observed pupil-teacher ratio

Box 2. The Service Delivery Indicators (SDI) Program

A significant share of public spending on education is transformed to produce good outcomes at schools. Understanding what takes place at these frontline service provision centers is the starting point in establishing where the relationship between public expenditure and outcomes is weak within the service delivery chain. Knowing whether spending is translating into inputs that teachers have to work with (e.g. textbooks in schools), or how much work effort is exerted by teachers (e.g. how likely are they to come to work), and their competency would reveal the weak links in the service delivery chain. Reliable and complete information on these measures is lacking, in general.

To date, there is no robust, standardized set of indicators to measure the quality of services as experienced by the citizen in Africa. Existing indicators tend to be fragmented and focus either on final outcomes or inputs, rather than on the underlying systems that help generate the outcomes or make use of the inputs. In fact, no set of indicators is available for measuring constraints associated with service delivery and the behavior of frontline providers, both of which have a direct impact on the quality of services that citizens are able to access. Without consistent and accurate information on the quality of services, it is difficult for citizens or politicians (the principal) to assess how service providers (the agent) are performing and to take corrective action.

The SDI provides a set of metrics to benchmark the performance of schools in Africa. The Indicators can be used to track progress within and across countries over time, and aim to enhance active monitoring of service delivery to increase public accountability and good governance. Ultimately, the goal of this effort is to help policymakers, citizens, service providers, donors, and other stakeholders enhance the quality of services and improve development outcomes.

The perspective adopted by the Indicators is that of citizens accessing a service. The Indicators can thus be viewed as a service delivery report card on education. However, instead of using citizens' perceptions to assess performance, the Indicators assemble objective and quantitative information from a survey of frontline service delivery units, using modules from the Public Expenditure Tracking Survey (PETS), Quantitative Service Delivery Survey (QSDS), and Staff Absence Survey (SAS).

The literature points to the importance of the functioning of schools and more generally, the quality of service delivery. The service delivery literature is, however, clear that, conditional on providers being appropriately skilled and exerting the necessary effort, increased resource flows for health can indeed have beneficial education outcomes.

The SDI initiative is a partnership of the World Bank, the African Economic Research Consortium (AERC), and the African Development Bank to develop and institutionalize the collection of a set of indicators that would gauge the quality of service delivery within and across countries and over time. The ultimate goal is to sharply increase accountability for service delivery across Africa, by offering important advocacy tools for citizens, governments, and donors alike; to work toward the end goal of achieving rapid improvements in the responsiveness and effectiveness of service delivery.

More information on the SDI survey instruments and data, and more generally on the SDI initiative can be found at: www.SDIndicators.org and www.worldbank.org/sdi, or by contacting sdi@worldbank.org.

II. METHODOLOGY AND IMPLEMENTATION

The sample of the Tanzania SDI is given in Table 5. Overall, 400 primary schools were visited, 2,197 standard three, four, and five teachers were assessed on English, mathematics, and pedagogy, 3,692 teachers of all grades have been followed for absence rate (not shown in Table 5). Also, although learning outcomes are not part of the indicators, 4,041 standard four pupils have been assessed on language (English/Kiswahili), mathematics, and non-verbal reasoning.[4] It is crucial that the indicators be correlated with pupil learning outcome because the SDI is geared towards capturing the drivers of learning outcomes at the school level.

Table 5. Tanzania's Education SDI Sample

	Schools				Teachers			Standard 4 Pupils		
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(1)	(2)	(3)
Total Sample	**400**				**2,197**			**4,041**		
Stratum										
Dar es Salaam	47	11.7	2.4	2.3	349	15.9	4.4	474	11.8	6.0
Other Urban	58	14.5	10.2	13.5	378	17.2	17.4	583	14.4	15.1
Rural	221	55.3	67.4	63.9	1,077	49.0	59.0	2,228	55.1	57.5
EQUIP-T	74	18.5	20.0	20.3	393	17.9	19.2	756	18.7	21.4
Location										
All Rural	287	71.7	85.1	79.2	1,427	65.0	74.7	2,901	71.8	75.8
All Urban	113	28.3	24.9	20.8	770	35.0	25.3	1,140	28.2	24.2

Notes: Each unit of analysis i.e. schools, teachers, or pupils has its own specific weights which are where relevant. Columns' definitions are as follows: (1) is sample size; (2) is share of sample; (3) is the weighted share; (4) is the distribution in the actual sample frame or universe.

The Tanzania SDI is representative of primary schools at the national level. It is also representative of Dar es Salaam, the main city, as well as other urban areas, and rural Tanzania. Because of a large DFID program called EQUIP-T being implemented in four regions (Dodoma, Kigoma, Tabora and Shinyanga), it was also decided to make these four regions as a whole a stratum. Therefore, the SDI can report statistics on the EQUIP-T regions as well. The sampling strategy is fully explained in the annex of this report. It is noteworthy that each entity has its own weight. Weights for schools are therefore different from weights from pupils or teachers. For the latter weights even differ for the analysis of absenteeism or the knowledge content analysis. The difference in weights comes from the fact that for each unit of analysis a sample needs to be drawn.

It is also noteworthy that the Tanzania Education SDI, unlike all other SDI countries, only reports on public schools. This is due to the near non-existence of private schools at the primary level. Indeed there are very few private schools in the sample frame and this is also confirmed by independent household surveys. As a matter of fact according to the 2012/13 National Panel Survey only two percent of the children attending primary school are in the private sector.

The education service indicators build on a growing body of literature on measuring the performance of schools and, specifically, of teachers. The Indicators provide a snapshot of the learning environment and key resources which need to be in place for pupils to learn. As the

[4] Following the Tanzania 2010 SDI results, it was decided for this round to add a Kiswahili test for the pupils. In each school roughly three pupils were tested in Kiswahili and seven in English (to maintain comparability with other countries and Tanzania's previous round). Overall 1,200 Standard 4 pupils were tested in Kiswahili and 2,841 sat for the English test.

expenditure on teachers represents, by far, the largest share of education spending in developing countries, and, as several recent studies convincingly demonstrate how changes in teacher behavior can improve learning achievement, a strong focus is placed on the knowledge, skills and effort of teachers. Annex B presents the Service Delivery Indicators in education and a short definition of each indicator. The survey instrument consists of the six modules as shown in Table 6.

A minimum requirement for learning is that the teachers are not absent from the school and spend time in the class rather than somewhere else. As shown in the definitions table in Annex B, the first two indicators, *School absence rate* and *Classroom absence rate*, are direct measures of the extent to which this is the case. While having teachers in the class is a necessary condition, however, it is not sufficient for learning to take place. Teachers need to be involved in teaching and need to have at least a minimum level of knowledge of the subjects they are teaching and skills to transform their knowledge into meaningful teaching. The indicator *Minimum knowledge* measures to what extent these skills exist across schools and the indicator Time spent teaching per day measures the extent to which teachers are exerting effort to enable learning.

Four of the indicators deal with the environment; i.e., the school. The indicator *Minimum equipment availability* assesses if necessary materials such as blackboard, chalk, pencils, and paper are in place to support the teaching activities. The indicator *Minimum infrastructure availability* measures whether functioning sanitation facilities exist and if there is at least minimum light in the classroom so that pupils can read and study. Finally, the indicators *Observed pupil-teacher ratio* and *Share of pupils with textbooks* measure the average number of pupils per teacher in grade four classrooms and the number of mathematics and language books at their disposal. Below we discuss each indicator in more depth and provide additional information on how they are derived, how they should be interpreted, and what they imply.

Table 6. Education SDI survey instrument

Module	Description
Module 1: School Information	Administered to the head of the school to collect information about school type, facilities, school governance, pupil numbers and school hours. Includes direct observations of school infrastructure by enumerators.
Module 2a: Teacher Absence & Info	Administered to head teacher and individual teachers to obtain a list of all school teachers, to measure teacher absence and to collect information about teacher characteristics.
Module 2b: Teacher Absence & Info	Unannounced visit to the school to assess absence rate.
Module 3: School Finances	Administered to the head teacher to collect information about school finances.
Module 4: Classroom Observation	An observation module to assess teaching activities and classroom conditions.
Module 5: Pupil Assessment	A test of pupils to have a measure of pupil learning outcomes in mathematics and language in grade four.
Module 6: Teacher Assessment	A test of teachers covering mathematics and language subject knowledge and teaching skills.

III. RESULTS

The indicators *Minimum equipment availability*, *Minimum infrastructure availability*, *Observed pupil-teacher ratio*, and *Share of pupils with textbook*[5] are all constructed using data collected through visual inspections of a Standard four classroom and the school premises in each primary school. Below we discuss each indicator in some more detail. Table 7 summarizes the findings.

A. Availability of inputs at the school

Minimum equipment availability

Methodological Note

Minimum equipment availability is a binary indicator capturing the availability of: (i) functioning blackboard and chalk and (ii) pens, pencils and exercise books in 4th grade classrooms. In one randomly selected 4th grade classroom in the school the enumerator assessed if there was a functioning blackboard by looking at whether text written on the blackboard could be read at the front and back of the classroom, and whether there was chalk available to write on the blackboard. We considered that the classroom met the minimum requirement of pens, pencils and exercise books if both the share of pupils with pen or pencils and the share of pupils with exercise books were above 90%.

Of the four indicators, *Minimum resource availability* appears less of a constraint. In terms of the availability of teaching resources, only six out of 10 (61.1 percent) Tanzanian primary schools seem to possess the minimum required. The main constraint is a functioning blackboard and the lack of light for pupils to be able to read the blackboard. Indeed, as shown in Table 7, all sub-indicators are close to 100 percent except "sufficient contrast to read board". In more than one out of four schools the standard four classroom was judged by the enumerator as not having enough contrast to allow proper reading from a distance. Schools in the EQUIP-T regions and rural schools in general are the hardest hit. Although lack of teaching equipment does not appear to be a binding constraint for providing high quality teaching in most Tanzanian primary schools, lack of light is a concern that needs to be addressed. As shown by Mott et al. (2012) lighting quality in the classroom may significantly impact learning process and outcomes.

[5] This indicator is used in lieu of *Pupils per textbook* which is the traditional indicator. The primary reason for this change is that this indicator is not defined in a classroom without textbook which proved to be a regular occurrence in previous SDI countries.

Minimum infrastructure availability

Methodological Note

Minimum infrastructure availability is a binary indicator capturing the availability of: (i) functioning toilets and (ii) classroom visibility. Functioning toilets is defined as whether toilets were functioning, accessible, clean and private (enclosed and with gender separation) as verified by an enumerator. To verify classroom visibility we randomly selected one 4th grade classroom in which the enumerator placed a printout on the board and checked whether it was possible to read the printout from the back of the classroom.

Share of pupils with textbooks reflects the typical ratio in pupil to textbooks in a 4th grade classroom. It is measured as the number of pupils with the relevant textbooks (language or mathematics conditional on which randomly selected class is observed) in one randomly selected 4th grade class and divided by the number of pupils in that classroom.

Observed pupil-teacher ratio reflects the typical ratio of pupils to teachers in a 4th grade classroom. It is measured as the number of pupils in one randomly selected 4th grade class at the school.

Table 7 reports the means for each sub-indicator for *Minimum infrastructure availability*. For a fuller breakdown of these results, see Table C 1. There is near universal access to toilets in Tanzania's primary schools and almost all of them are accessible and clean.[6] Because in many of the schools teachers do not have separate toilets and need to share in pupils' toilets, it is considered that pupils enjoy no privacy in those toilets. As a result, fewer than three out of five (57 percent) of the schools are considered as having private toilets. It is especially in rural areas as well as EQUIP-T regions that this phenomenon is widespread. On the last sub-indicator - "visibility" – it is again in rural areas and EQUIP-T regions that pupils will have more difficulty reading what is written on the board.

Overall, Tanzanian schools score poorly on *Minimum infrastructure availability* with only 41 percent of them meeting the standard. The major constraint for infrastructure is the privacy of the toilets because they are shared with teachers especially in rural areas. Again visibility in the classrooms is an issue for rural schools.

[6] For this indicator, there were a number of schools that do not, in fact, have information on cleanliness and accessibility. The specific schools with missing information were those where teachers and pupils used the same toilets. For such schools, a skip was inadvertently included in the Kiswahili version of the questionnaire. We considered that those schools did not meet the privacy criterion for pupils. The toilets were, however, still considered accessible and clean by default. The infrastructure indicator, therefore, is an overestimate of the true state of infrastructure in Tanzania's primary schools.

Table 7. At the School, auxiliary information

(Percent)	Tanzania	Dar es Salaam	Other Urban	Rural	EQUIP-T	All Rural	All Urban
Minimum resource availability	**61.4**	**83.2**	**80.1**	**58.3**	**58.9**	**58.3**	**80.4**
Share of pupils with pencil	95.8	96.9	96.7	95.7	97.0	95.7	96.7
Share of pupils with paper	96.3	96.8	97.3	96.2	97.5	96.2	97.2
Have a board	98.4	99.1	97.2	98.5	94.4	98.5	97.5
Have chalk	97.0	100.0	96.2	97.1	100.0	97.1	96.8
Sufficient contrast to read board	73.9	93.1	88.4	71.4	66.2	71.4	89.0
Minimum infrastructure availability	***40.4***	***67.0***	***61.4***	***36.8***	***34.3***	***36.9***	***62.3***
Visibility (by enumerator)	75.8	94.0	89.4	73.4	71.8	73.5	90.1
Toilet clean	92.0	95.3	96.1	91.4	90.7	91.4	95.9
Toilet private	56.2	83.3	79.7	52.2	53.4	52.3	80.6
Toilet accessible	96.6	94.4	92.9	97.2	98.0	97.2	93.1
Observed pupil-teacher ratio	43.5	69.8	58.9	40.7	46.1	40.8	60.4
Share of pupils with textbooks	25.3	31.0	14.1	26.7	22.2	26.7	16.7
Mathematics textbook	24.6	28.4	12.2	26.2	19.8	26.3	14.7
English textbook	26.3	34.9	16.8	27.4	25.9	27.4	19.7

Source: Tanzania SDI 2014 and author's calculations

Roughly only one out four pupils had access to an English or mathematics textbook in a typical standard four classroom. Table 7 provides the statistics on the share of pupils who had or were sharing a textbook broken down by subject area (English and mathematics). First, it is important to state that in 42 percent of the schools none of the pupils had a textbook during the lesson. In the schools in which textbooks are available every three children would have to share two textbooks. There is no significant difference across subjects with mathematics books being as prevalent as English books. There seems also to be little variation across strata with maybe the exception of urban areas other than Dar which have a slightly lower share of pupils with a textbook during class. This lack of books in the classrooms is observed even after a large distribution of textbooks took place recently in Tanzania and for which schools in the SDI survey acknowledge being beneficiaries. The reason is that often schools may have the textbooks but decide not provide them to pupils. In a recent study in Sierra Leone, Sabarwal et al. (2014) concludes that schools that have high uncertainty with respect to future transfers are more likely to store a proportion of current transfers (textbooks) with a view towards smoothing "consumption". It is not clear whether this is what happens in Tanzanian schools but the phenomenon is worth further investigation to make sure books are fully and efficiently used by schools and pupils.

The *Observed pupil-teacher ratio* stand at 43 in the classrooms observed. The ratio is slightly below Tanzania's recommended benchmark of 45:1. There are large and significant differences across strata. With close to 70 pupils, Dar es Salaam's average standard four classroom has 27 more pupils than the typical standard four classroom in the nation or 60 percent more pupils.[7] Other urban schools also experience overcrowding of classrooms albeit to a lower extent. Interestingly, it is rural schools which respect the norm in terms of class size and this may be due to the recent acceleration of urbanization in Tanzania as shown by the recent economic update.

[7] Note that this is fairly different from Uwezo's 2012 results which found a significantly lower PTR in Dar compared to other regions. However, this is consistent with administrative data on school enrollments and should hold if pupil's absenteeism is similar across regions. Note also that Dar schools have on average more classrooms but the difference is not large enough to make their class sizes smaller.

B. Teachers effort

There are three indicators designed to capture the effort teachers put into their job. These indicators are (i) *School absence rate*, (ii) *Classroom absence rate*, and (iii) *Time spent teaching per day.* The rationale behind these indicators is that the low levels of accountability and weakened incentives observed in many countries especially in sub-Sahara Africa have led to an upsurge of no-show for teachers. A classroom with no teacher is an environment where no learning is taking place. The indicators are computed across strata to capture the variations in these important dimensions in the country.

School absence rate

Methodological Note

School absence rate is measured as the share of teachers who are absent from school at the time of an unannounced visit. It is measured in the following way: During the first announced visit, a maximum of ten teachers are randomly selected from the list of all teachers (excludes volunteer and part time teachers) who are on the school roster. The whereabouts of these ten teachers are then verified in the second, unannounced, visit. Teachers found anywhere on the school premises are marked as present.

As shown in Table 8, the school absence rate is relatively low with one out of seven (14.3 percent) not present at school at the time of the surprise visit.[8] Teachers in rural schools are equally likely to be absent from school than their urban colleagues. Absence rate is lowest in the EQUIP-T regions with "only" 8.3 percent of the teachers absent at any given time. Dar es Salaam's public school teachers are more than twice as likely to be absent from school than teachers in EQUIP-T regions.

Table 8. School absence rate and Classroom absence rate

	School absence rate				Classroom absence rate			
	Percent	Robust Std. Err.	[95% Conf. Interval]		Percent	Robust Std. Err.	[95% Conf. Interval]	
Tanzania	14.4	0.010	12.3	16.4	46.7	1.2	44.3	49.1
Dar es Salaam	17.8	2.3	13.4	22.3	42.9	3.5	36.0	49.8
Other Urban	13.7	1.7	10.3	17.0	46.7	2.3	42.3	51.1
Rural	14.3	1.3	11.7	16.9	47.0	1.5	44.1	50.0
EQUIP-T	8.3	1.5	5.3	11.3	43.4	2.6	38.3	48.5
All Rural	14.3	1.3	11.7	16.9	47.0	1.5	44.1	49.9
All Urban	14.6	1.4	11.8	17.4	45.8	1.9	42.0	49.6

Source: Tanzania SDI 2014 and author's calculations

[8] The majority of the surprise visits took place during the morning with roughly 70 percent of the enumerators arriving before 12am (the mode of arrival is between 9-10 am). The surprise visit lasted 45 minutes on average.

Classroom absence rate

Methodological Note

Classroom absence rate is measured as the share of teachers not in the classroom at the time of an unannounced visit. The indicator is constructed in the same way as the school absence rate indicator, with the exception that the numerator now is the number of teachers who are either absent from school, or present at school but absent from the classroom.

Even when in school, teachers may not necessarily be in the classroom teaching. To capture this new dimension, the indicator *Absence from class* is used. A teacher is considered absent from classroom if she is either not on the school premises or when in school, she cannot be located inside a classroom.[9]

Close to two out of five (37 percent) of the teachers found in school are not in the classroom teaching. This brings the absenteeism from classroom to 46.5 percent nationally. This simply means that at any point in time, almost half of Tanzanian primary teachers are outside the classroom and are thus not teaching. As expenditure on teachers represents by far the largest share of education spending in developing countries including Tanzania, this very high absence from classroom clearly constitutes an important waste of time and resources with half of the time of teachers not utilized interacting with their pupils. Interestingly, absence from classroom does not vary much across regions.

Table 9. "Quartiles" of absence rates

	"Quartiles" of School absence rate				"Quartiles" of Classroom absence rate			
	20% or less	*20 to 40%*	*40 to 60%*	*60+% absent*	*20% or less*	*20 to 40%*	*40 to 60%*	*60+% absent*
Tanzania	64.8	25.9	6.4	2.9	7.7	22.5	42.8	27.0
Dar es Salaam	40.7	46.0	13.4	0.0	8.8	19.4	40.2	31.6
Other Urban	63.8	28.4	4.6	3.2	5.3	22.6	48.1	24.0
Rural	65.6	25.0	6.4	3.0	8.0	22.6	42.1	27.3
EQUIP-T	81.8	12.5	4.2	1.5	10.3	25.3	39.2	25.2
Rural	65.5	25.1	6.4	3.0	8.0	22.6	42.2	27.2
Urban	60.6	30.5	6.2	2.7	5.9	22.2	46.2	25.7

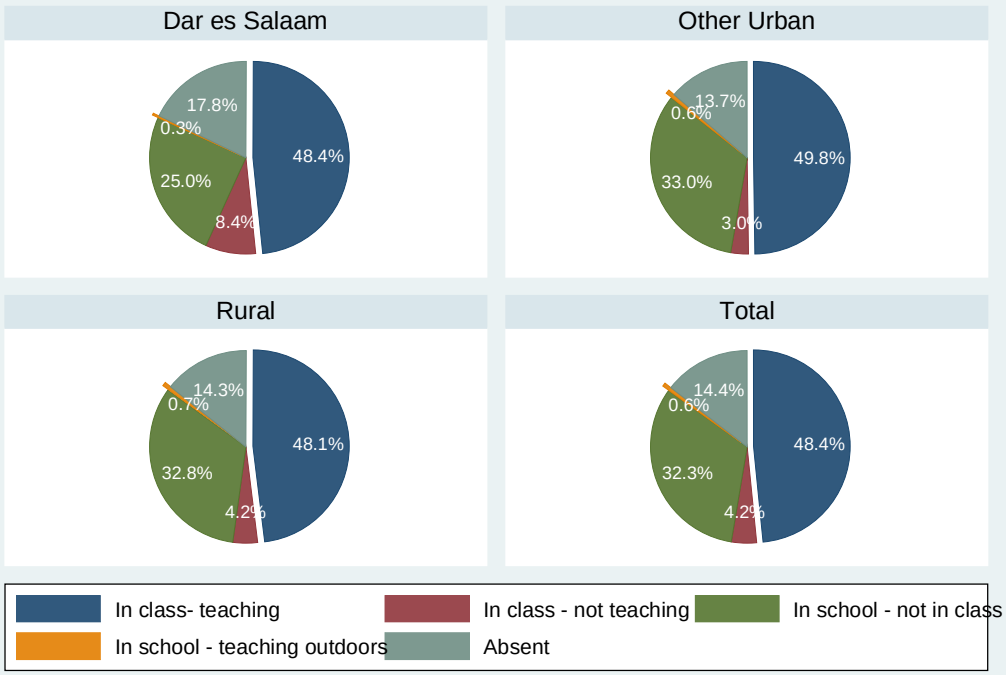
Source: Tanzania SDI 2014 and author's calculations

Table 9 provides information on the distribution of absence by looking at the "quartiles". The salience of classroom absence and its contrast compared to school absence appears even more strikingly. Indeed, a large majority (64.8 percent) of schools record school absence rates below 20 percent. However, fewer than one out of 10 schools (7.7 percent) have a classroom absence in the first quartile. In most schools (42.8 percent), four to six teachers out of 10 will not be in the classrooms and

[9] A small number of teachers are found teaching outside, and these are marked as present for the purposes of the indicator.

in a quarter of the schools more than 60 percent of the teachers are not in the classrooms. The distribution of classroom absence is fairly similar across strata although EQUIP-T regions stand out as shown in Figure C 1. Classroom absence is most likely a school leadership and management issue because the majority of the teachers are in fact in the school. The recent EQUIP-T (2015) report echoes this concern and shows that head teachers have a weak understanding and implementation of their role and responsibilities.

Figure 2. Teachers' whereabouts during unannounced visit



Source: Tanzania SDI 2014

Where were the teachers at the time of the unannounced visit? Figure 2 provides the answer to that question across strata and offers yet another perspective on absence rates by showing teachers' whereabouts during the surprise visit. Confirming Table 8 and Table 9 about half of the teachers are teaching, some of them (0.6 percent nationally) outside the classroom. Although classroom absence rate is at 46.7 percent, only 49 percent of the teachers were actually teaching. Indeed, a non-negligible share (4.2 percent) of teachers was in class but attending to other matters than teaching. The "In class but not teaching" phenomenon is more prevalent in Dar es Salaam with 8.4 percent (almost one out of 10) of teachers in that mode – double the national average.

Finally, Table 10 shows absence rates as related to few select teachers' characteristics, such as gender, place of birth, and position in school. As for absence across strata, there are no clear correlates of school or classroom absence. There is no statistically significant gender difference with female teachers equally likely to be absent than male teachers. Although teachers who were born in the district in which they work seem more likely to be absent from school, again that difference is not statistically significant and they are no different than those who work in a district they were not born

into. The only significant difference is for classroom absence between head teachers and regular teachers. This may be explained by the administrative and other duties that head teachers have on top of their teaching load when they do teach. It is, therefore, not surprising that head teachers are found less often in the classroom. Interestingly, and unlike many other SDI countries, head teachers are less likely to be absent from school than regular teachers even though the difference is not significant.

Table 10. School absence rate and Classroom absence rate by gender and birth place

	School absence rate				Classroom absence rate			
	Percent	Robust Std. Err.	[95% Conf. Interval]		Percent	Robust Std. Err.	[95% Conf. Interval]	
Tanzania	14.4	1.0	12.3	16.4	46.7	1.2	44.3	49.1
Male	13.5	1.2	11.1	16.0	46.6	1.8	43.1	50.2
Female	15.3	1.3	12.7	17.8	46.7	1.4	43.9	49.6
Head teacher	12.7	3.0	6.8	18.6	59.2	4.1	51.2	67.3
Other teacher	14.5	0.9	12.6	16.4	45.4	1.2	43.0	47.7
Born in district	16.5	2.9	10.8	22.2	47.9	2.6	42.8	53.0
Not born in district	13.4	1.0	11.4	15.3	45.8	1.3	43.3	48.4

Source: Tanzania SDI 2014 and author's calculations

Time spent teaching per day

Methodological Note

Time spent teaching per day reflects the typical time that teachers spend teaching on an average day. This indicator combines data from the staff roster module (used to measure absence rate), the classroom observation module, and reported teaching hours. The teaching time is adjusted for the time teachers are absent from the classroom, on average, and for the time the teacher teaches while in classroom based on classroom observations. While inside the classroom distinction is made between teaching and non-teaching activities.

Teaching is defined very broadly, including actively interacting with pupils, correcting or grading pupil's work, asking questions, testing, using the blackboard or having pupils working on a specific task, drilling or memorization. Non-teaching activities include working on private matters, maintaining discipline in class or doing nothing and thus leaving pupils not paying attention.

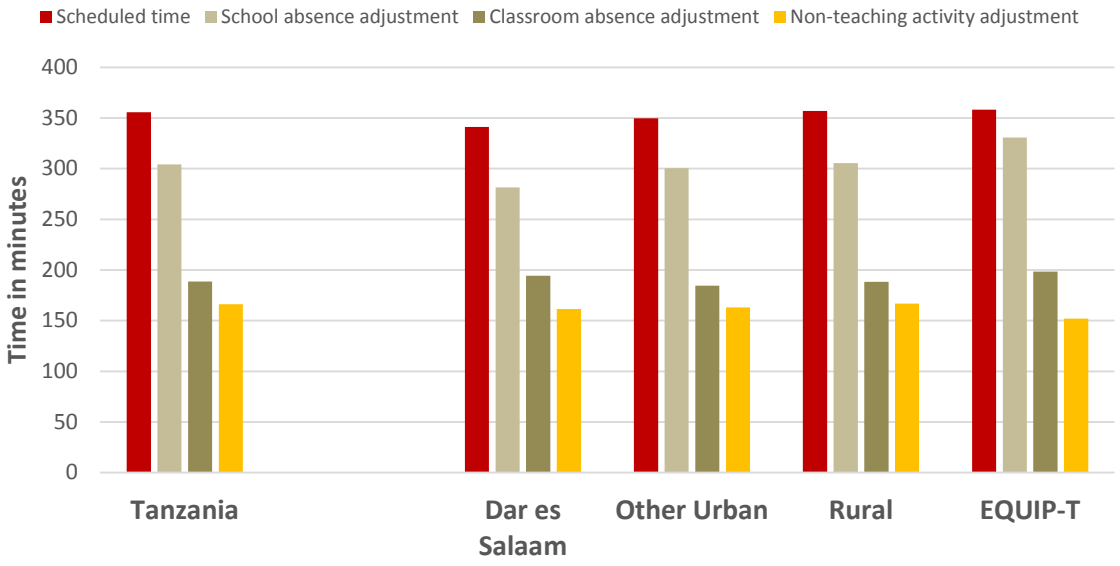
This indicator measures the amount of time a teacher spends teaching in a school during a normal day, which on average was 2 hours and 47 minutes in Tanzania for the 2013/2014 school year (Table C 2). That is, teachers taught only about half of the scheduled time (which is 5 hours and 56 minutes accounting for break times). Several intermediate inputs feed into the calculation of this indicator and are reported in Figure 3.

The first step was to begin by recording the scheduled time of a teaching day from school records, which was 5 hours and 56 minutes on average. Then we multiplied this number by the proportion of teachers absent from school. The idea was that if 10 teachers were supposed to teach 5 hours and 56

minutes per day, but one of them was nowhere to be found in school, then scheduled teaching time was reduced to 5 hours and 4 minutes (5 hours and 56 minutes x 0.856). The second step was to adjust with classroom absence (i.e., teachers may have been at the school, but they were not in the classroom teaching). Given the importance of classroom absence, the scheduled teaching time was reduced further to 3 hours and 10 minutes (5 hours and 56 minutes x 0.533[10]).

The last step consisted of taking into account and removing the time lost by teachers on non-teaching activities while in the classroom. Indeed, even when in the class, teachers may not necessarily be teaching. The percentage of the lesson lost to non-teaching activities was measured through observation of a standard four lesson.[11] As reported in Table C 2, roughly 12 percent of a typical lesson was lost due to non-teaching activities.[12] To take this into account, we multiplied our measure by the proportion of a typical lesson that was spent on teaching. In the example, the teaching time of 3 hours and 10 minutes fell again to a low 2 hours and 47 minutes (3 hours and 10 minutes x 0.88).

Figure 3. From official scheduled time to effective teaching time



There was no major difference in learning time across strata. Pupils in the EQUIP-T regions seemed to receive slightly less than 20 minutes of learning compared to their urban peers but again this difference was not statistically significant. It is noteworthy that EQUIP-T pupils spent less time (albeit not significantly so) interacting with their teachers despite the fact that they (1) had the highest

[10] The numbers 0.857 = (1- 0.144) and 0.533 = (1-0.467) represent the share of teachers in school and the classroom respectively.

[11] This is most likely an upper bound on the time devoted to teaching during a lesson, since presumably a teacher is more likely to teach when under direct observation (i.e. Hawthorne effects will bias the estimate upward).

[12] During the observation, enumerators first had to judge whether the teacher was teaching or not. If they judged the teacher to be teaching, they were supposed to indicate how much time the teacher spent on any of the following teaching activities: teacher interacts with all children as a group; teacher interacts with small group of children; teacher interacts with children one on one; teacher reads or lectures to the pupils; teacher supervises pupil(s) writing on the board; teacher leads kinesthetic group learning activity; teacher writing on blackboard; teacher listening to pupils recite/read; teacher waiting for pupils to complete task; teacher testing pupils in class; teacher maintaining discipline in class; teacher doing paperwork.

scheduled teaching time, and (2) enjoyed a slightly lower teacher absence rate. The main reason was that they also had the highest time loss (20 percent, see Table C 2) on non-teaching activities when the teacher was in the classroom.

Finally, Table 11 (and Figure C 2) provides information on a complementary measure of effort – the share of classrooms with pupils but no teacher; i.e., orphaned classrooms. This was measured by inspecting the school premises, counting the number of classrooms with pupils, and recording whether a teacher was present in the classroom or not. The share of orphaned classrooms was then calculated by dividing the number of classrooms with pupils but no teacher by the total number of classrooms that contained pupils. In total, about one out of three classrooms (35 percent) were orphaned and again there was little variation across Tanzania for this measure of effort.

Table 11. Orphaned classrooms (No teacher but pupils were present)

	Total # Classrooms	# Classrooms with Pupils	# Orphan Classrooms	Percent	Std. Err.	[95% Conf. Interval]	
				Share Orphan Classrooms			
Tanzania	7.8	7.0	2.5	35.3	1.8	31.8	38.8
Dar es Salaam	15.1	13.0	4.4	33.8	4.4	25.1	42.4
Other Urban	9.6	8.6	3.0	34.5	4.7	25.2	43.9
Rural	7.3	6.6	2.4	35.6	2.0	31.6	39.5
EQUIP-T	7.1	6.7	2.2	32.6	4.3	24.3	41.0
All Rural	7.3	6.7	2.4	35.5	2.0	31.6	39.5
All Urban	10.5	9.3	3.2	34.4	3.8	26.9	41.9

Source: Tanzania SDI 2014 and author's calculations

Correlations between teacher effort and school inputs

One might expect that better infrastructure would be associated with more teacher effort – at least poor quality infrastructure is often named by teachers as a reason for low motivation. Looking at the SDI data, however, there is little evidence that school resources are correlated with teacher effort. Examining the correlations between *School absence rate* and *Classroom absence rate* and the various infrastructure indicators, no consistent picture emerges. While school absence is negatively correlated with the sub-indicator *Share of pupils with pencils*, it is not correlated with the broader indicator *Minimum equipment availability*. School absence seems independent of all other traditional input indicators such as *Minimum infrastructure availability* or *Share of pupils with textbooks*. In the same vein, classroom absence is negatively correlated with the *Share of pupils with textbook*s and independent of all other indicators included in Table 12. *Time spent teaching per day*, on the other hand, is not correlated with any indicator or sub-indicator. Teacher's effort in Tanzania, therefore, does not seem to be influenced by a school's physical resources. An alternative explanation for this lack of correlation might be the observed lack of variation in most of the indicators in Tanzania (at least on average) across strata.

To summarize this section's analysis, one observes relatively low school absence; i.e., by and large, teachers showed up at school, but once they were on the school premises, many of them did not spent much time in the classrooms teaching their pupils. In a large share of classrooms, pupils were by themselves while the teacher was engaged in activities not related to teaching. By

this process, Tanzanian pupils lost more than half of the teaching time they were supposed to receive and were able to interact with their teachers only 2 hours and 47 minutes per day in lieu of the official 5 hours and 55 minutes.

Table 12. Correlates of teachers' effort

	School absence			Classroom absence			Time spent teaching		
	Coef	R2	N	Coef	R2	N	Coef	R2	N
Infrastructure	0.00408	0.000	400	-0.00146	0.000	400	-10.47	0.003	399
	(0.0166)			(0.0232)			(9.822)		
Teaching Equipment	-0.00666	0.000	400	-0.000354	0.000	400	2.996	0.000	399
	(0.0177)			(0.0247)			(10.49)		
Pupils with pencils (%)	**-0.152****	0.003	400	-0.150	0.006	397	56.35	0.004	396
	(0.0714)			(0.100)			(42.59)		
Pupils with notebooks (%)	-0.127	0.002	400	-0.135	0.003	397	67.36	0.005	396
	(0.0831)			(0.116)			(49.41)		
Class has board	0.00425	0.011	397	-0.0164	0.000	400	-6.341	0.000	399
	(0.0745)			(0.104)			(44.11)		
Pupils with textbooks (%)	-0.0373	0.004	399	**-0.0645***	0.007	399	12.45	0.001	398
	(0.0280)			(0.0391)			(16.75)		
Pupils in class	-0.00023	0.002	400	5.75e-05	0.000	400	-0.177	0.003	399
	(0.00026)			(0.00037)			(0.156)		

Note: Standard errors in parentheses. Levels of significance: *** p<0.01, ** p<0.05, * p<0.1. Source: Tanzania SDI 2014 and author's calculations.

C. Teachers competence

Even if teachers show up to school and spend the allocated time in the classroom engaging in teaching activities with their pupils as expected, they need to have a fairly good command of the subject they teach as well as the required pedagogical skills to effectively pass that knowledge unto their pupils. This section discusses the indicator used to capture teachers' knowledge and capacity to teach which is dubbed *Minimum knowledge*.

Minimum knowledge

Methodological Note

Minimum knowledge is measured as the percentage of teachers who can master the curriculum they taught. It is based on a mathematics and language tests covering the primary curriculum administered at the school level and is calculated as the percentage of teachers who score more than 80% on the language and mathematics portion of the test. The test is given to all mathematics or language teachers that taught 3rd grade last year or 4th grade in the year the survey was conducted.

The share of teachers with minimum subject content knowledge is calculated on the basis of a custom-designed teacher test administered to the standard four mathematics and English teachers of the 2013 and 2014 pupil cohorts. The objective of the teacher test is to examine whether teachers

have the basic reading, writing, and arithmetic skills that lower primary pupils need to have in order to progress further with their education. Teachers are considered capable of teaching if they score at least 80 percent on the tests, which is interpreted as the *minimum* knowledge required for the teacher to be effective.

In addition, the test also examines the extent to which teachers demonstrate mastery of subject content skills that are above the level they are teaching at and mastery of pedagogic skills. Out of courtesy to teachers the test was designed as a marking exercise, in which teachers had to mark and correct a hypothetical pupil's exam. The English (mathematics) test was administered to teachers teaching English (mathematics), although they might teach other subjects as well. The test was validated against the Tanzanian primary curriculum as well as 12 other Sub-Saharan curricula.[13]

The minimum knowledge indicator is calculated as the percentage of teachers who score more than 80 percent on the English and mathematics test. The test also contains more advanced questions in both subjects as well as a pedagogy section.

Table 13. Teachers' tests performance (English and mathematics combined)

(Percent)	Tanzania	Dar es Salaam	Other Urban	Rural	EQUIP-T	All Rural	All Urban
Minimum knowledge (English & mathematics) *(% of teachers)*							
Cut-off point*: 80%	***21.5***	***26.2***	***22.3***	***21.0***	***19.5***	***21.0***	***23.2***
Min. Knowledge (English): 80%	1.1	2.1	0.5	1.1	0.3	1.1	0.8
Min. Knowledge (Maths): 80%	25.5	33.2	20.8	26.2	21.0	26.2	22.9
Average Scores (%)							
English & mathematics	58.9	65.1	60.4	58.3	57.4	58.3	61.3
English	41.9	52.2	44.7	40.8	40.6	40.8	45.9
Mathematics	63.1	69.1	64.6	62.5	59.3	62.5	65.4
Pedagogy	35.9	42.3	37.1	35.2	36.5	35.2	38.0

Note: (*) is the agreed cut-off point to estimate minimum knowledge for the SDI. Weighted means using sampling weight. Results based on observations from 2,150 teachers in 400 schools. 1,200 teachers either teach English or both English and mathematics and 1,483 teachers who teach either mathematics or both English and mathematics.

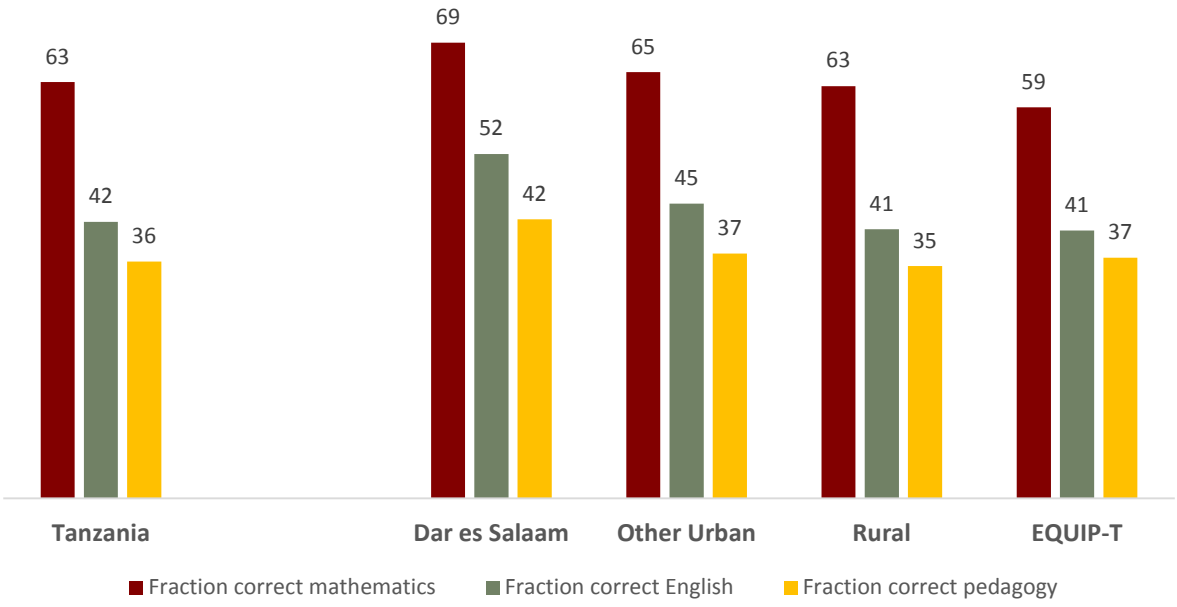
Content knowledge among Tanzanian teachers seems extremely low. As a matter of fact, only one out of five (21.5 percent) teachers score more than 80 percent on the combined mathematics and English test (Table 13). A familiar feature by now, there is no significant difference across strata even though a higher share (26 percent) of teachers in Dar es Salaam manages to score 80 percent or more, the difference with other strata is small and not significant. Overall however, all display equally disappointing levels of content knowledge.

Table 13 details the average score on the various parts of the assessment. Table 13 shows the extremely low scores on the English section -- a shocking 1.1 percent of teachers were above the 80 percent cut-off and this partly accounts for the overall low scores. Although mathematics teachers fared better, only one out of three (25.5 percent) were above the 80 percent cut-off.

[13] See "Teaching Standards and Curriculum Review", prepared as background document for the SDI by David Johnson, Andrew Cunningham and Rachel Dowling. The countries included for the review were: Botswana, Ethiopia, Gambia, Kenya, Magadascar, Mauritius, Namibia, Nigeria, Rwanda, Seychelles, South Africa, Tanzania and Uganda.

Table 13 and Figure 4 show the scores in the tests. Unlike the *Minimum knowledge* indicator, which captures the share of teachers who score above a threshold, the scores simply capture the proportion of questions teachers – taken as a whole- were able to answer as a share of the total number of questions. Taking mathematics and English together, teachers found roughly three out of five (58.9 percent) questions on the whole test. Clearly scores in mathematics are higher compared to English or pedagogy across all strata. Mathematics teachers correctly found two out of three questions (63.1 percent). As expected, teachers score poorly in English with an average score of 42.2 percent but they score even worse in pedagogy.

Figure 4. Teachers' average score on English, mathematics, and pedagogy tests



Tanzanian teachers, therefore, have a low pedagogic content knowledge which is critical for effective teaching. As a matter of fact, even when teachers have a deep knowledge of their subject matter, only a decent level of pedagogic knowledge content ensures that they can pass that knowledge onto their pupils. In Tanzania, there is room for progress for teachers to reach and stay at a reasonable level for both subject matter knowledge and pedagogic content knowledge.

Dar es Salaam teachers did significantly better than teachers in EQUIP-T regions across all subjects and did significantly better than rural teachers only in mathematics. The difference between Dar es Salaam teachers and other urban teachers was significant for none of the subject matters except for pedagogy, if one adopts a looser confidence level of 10 percent. The next sections carry out a more in-depth investigation of teachers' performance on each of the three tests (i.e., English, mathematics and pedagogy).

D. Test scores

Methodological Note

Test score is measured as the overall score of a language, mathematics and pedagogy tests covering the primary curriculum administered at the school level to all mathematics and language teachers that taught 3rd grade last year or 4th grade in the year the survey was conducted.

English

Table 14 presents the average score on the English test, as well as a detailed analysis of particular questions. The average score is 42 percent correct answers indicating that teachers do not even master half of the standard four curriculum. Nevertheless, this gives a much more positive picture than the *Minimum knowledge* indicator, whereby only 1.4 percent of the teachers scored more than 80 percent on the test. Even lowering the bar to a 70 percent mastery of the curriculum, only 4.9 percent of the teachers reached the bar.

Teachers scored an average close to 75 percent on the grammar assessment, which asked them to complete sentences with the correct conjunction, verb (active or passive voice and different tenses) or preposition. Four alternatives, including the correct one, were given for each sentence. Despite the good grammar scores, there were some serious gaps. For example, more than half (55 percent) of the teachers were not able to correct the sentence *"If you tidy up your room, you won't get candy"*, even though the correct alternative ("unless") was given (recall that teachers were asked to mark a hypothetical pupil's exam).

Table 14. Teachers' performance on English test by sub-section

(Percent)	Tanzania	Dar es Salaam	Other Urban	Rural	EQUIP-T	All Rural	All Urban
Minimum Knowledge[14]							
Cut-off point: 80%	***1.1***	***2.1***	***0.5***	***1.1***	***0.3***	***1.1***	***0.8***
Average scores							
English (complete test)	41.9	52.2	44.7	40.8	40.6	40.8	45.9
Grammar task	73.0	79.3	77.6	71.6	72.7	71.7	77.8
Cloze task	53.0	64.5	58.1	51.3	49.3	51.3	59.2
Composition task	21.3	32.9	21.9	20.6	20.8	20.6	23.8

Source: Tanzania SDI 2014 based on author's calculations.

Scores on the Cloze exercise which assesses vocabulary and text comprehension were somewhat lower (53 percent). The exercise consisted of a short story with certain words removed, and the teachers had to fill the gaps in a meaningful way. Again, some weaknesses emerged. While teachers were able to confirm that pupils had answered correctly, they struggled to correct wrong answers or complete sentences that the pupil had left blank. For example, 71 percent of teachers could not correct the sentence *"I **want** not go to school."*

[14] Note that the apparent inconsistency of having a minimum knowledge score of both mathematics and English greater than the minimum knowledge for only English or only mathematics comes from the fact that some teachers are evaluated on only one subject.

Teachers recorded their worst performance on the composition exercise with an average score of only 21 percent. They were tasked to correct a letter for grammar, punctuation, spelling, syntax and salutation. Sentences like *"I went to tell you that my new school is better the oldone"* were set for correction. Overall, the text to be corrected contained 21 errors and the teachers, on average, caught 4.6 mistakes. Only 25 percent of the teachers found and corrected seven or more mistakes in the standard four level six-line English paragraph.

Teachers in Dar es Salaam and other urban areas performed significantly better in grammar than teachers in EQUIP-T regions or rural areas. Dar es Salaam teachers actually significantly outperformed EQUIP-T and rural teachers in all sub-sections of the English test. When compared to other urban teachers, though, the difference in score was only statistically significant for the overall English score.

Mathematics

Table 15 presents the performance of teachers on the mathematics test, as well as a detailed analysis of particular questions. (For a full breakdown of the mathematics results, see Table C 3.) First, it is interesting to note that nationally about 1 in 100 teachers had a perfect score and 9 percent found and corrected 90 percent of the mathematics questions. Slightly more than one out of 4 (25.5 percent) of the teachers met the *Minimum knowledge* standard of 80 percent. There is clearly a lot of room for progress, yet this is undoubtedly a much more positive and encouraging picture compared to English.

The average score on the mathematics section was 63 percent correct answers and we observed a large and significant difference (19.2 percentage points) between scores on the lower and upper parts of the primary curriculum. This means that teachers were much more at ease with simple operations, such as adding two- or three-digit numbers, than with slightly more complex computations such as comparing fractions. This pattern holds true in all parts of the country even though standard four teachers are expected to master and be equally at ease with all parts of the curriculum. As for the English test, Dar es Salaam teachers outperformed teachers in EQUIP-T regions or rural areas but their scores were, on average, not statistically different from those of urban teachers. It was for only two questions that Dar teachers outperformed other urban school teachers; these are (i) computation of a square root and (ii) geometry 2D shapes (namely, finding the number of sides of a triangle).

Looking at the details of the test from Table 15, 14 percent of the teachers could not subtract two-digit numbers; the same proportion could not multiply two-digit numbers; one out of three could not add or subtract numbers with decimals; and roughly every other teacher could not perform division with fractions, solve a one variable equation, or interpret a Venn diagram. Almost three out of four teachers could not interpret data on a graph, which is upper primary level material, but should be mastered by teachers.

Table 15. Teachers' performance on mathematics assessment (and selected examples)

(Percent)	Tanzania	Dar es Salaam	Other Urban	Rural	EQUIP-T	All Rural	All Urban
Minimum Knowledge Cutoff point: 80%	***25.5***	***33.2***	***20.8***	***26.2***	***21.0***	***26.2***	***22.9***
Average scores							
Mathematics (complete test)	***63.1***	***69.1***	***64.6***	***62.5***	***59.3***	***62.5***	***65.4***
Lower primary	69.8	75.9	71.7	69.2	66.9	69.2	72.5
Upper primary	50.6	56.3	51.5	50.1	44.9	50.1	52.3
Adding two-digit numbers	97.0	99.1	97.3	96.8	95.1	96.8	97.6
Subtracting two-digit #	85.8	87.9	85.2	85.8	81.9	85.8	85.5
Adding three-digit numbers	85.3	91.2	89.5	84.1	86.2	84.1	90.0
Multiplying two-digit numbers	84.9	85.8	85.1	84.9	84.0	84.9	85.3
Adding decimals	64.0	63.9	66.5	63.4	62.3	63.5	66.0
Comparing fractions	81.3	85.1	86.8	80.0	81.2	80.0	86.6
Time (reading a clock)	47.5	55.8	55.3	45.4	40.7	45.3	55.7
Interpreting a Venn diagram	48.9	58.4	49.8	48.3	42.6	48.3	51.2
Interpreting data on a graph	27.5	24.4	25.5	28.1	23.8	28.0	25.5
Square root (no remainder)	79.3	92.6	78.5	78.9	73.7	78.9	80.8
Subtraction of decimal #	66.5	74.5	66.6	66.1	62.7	66.2	67.8
Division of fractions	58.3	63.0	58.6	58.0	53.6	58.0	59.4
One-variable algebra	50.5	53.6	50.9	50.2	44.2	50.2	51.3

Source: Tanzania SDI 2014 based on author's calculations

Pedagogy

The overall score on the pedagogy section was 36 percent with little difference between basic and more advanced questions (Table 16).[15] On average, teachers only managed about one-third of the tasks in the pedagogic test. Although all teachers seem to lack pedagogical skills, the extent is more serious in rural schools or schools in EQUIP-T regions when compared to Dar es Salaam. Overall, teachers in urban schools are no different than their rural counterparts when it comes to pedagogic content knowledge: they are all equally ill-equipped. This is clearly illustrated by the "minimum knowledge" indicator whereby only one out of 1000 (0.1 percent) teacher scored more than 80 percent on the pedagogy test. Even lowering the standard to 50 percent, only one out of five teachers (Figure 5) would reach the bar.

The pedagogy test consisted of three sections designed to capture all the skills teachers would routinely be asked to apply when teaching.[16] The first section asked teachers to prepare a lesson plan about road accidents in Tanzania based on a simple information-giving text they had read. The average score on this task was the lowest at 18 percent. The second task asked teachers to assess and compare children's writing on the basis of two sample letters. The average score on this task was 58

[15] Actually, quite surprisingly, teachers scored better on advanced pedagogy compared to basic pedagogy.

[16] Because the aim is to measure pedagogical skills, not English, the text was written in Kiswahili, except in places where English was part of the exercise.

percent. The final task asked teachers to inspect test scores of 10 children, aggregate them and make some statements about patterns of learning. This task received a score of 22 percent.

The low scores on the pedagogy section combined with the performance on the curriculum content imply that teachers know little more than their pupils and that the little they know, they cannot teach adequately.

Table 16. Teachers' performance on pedagogy assessment (selected examples)

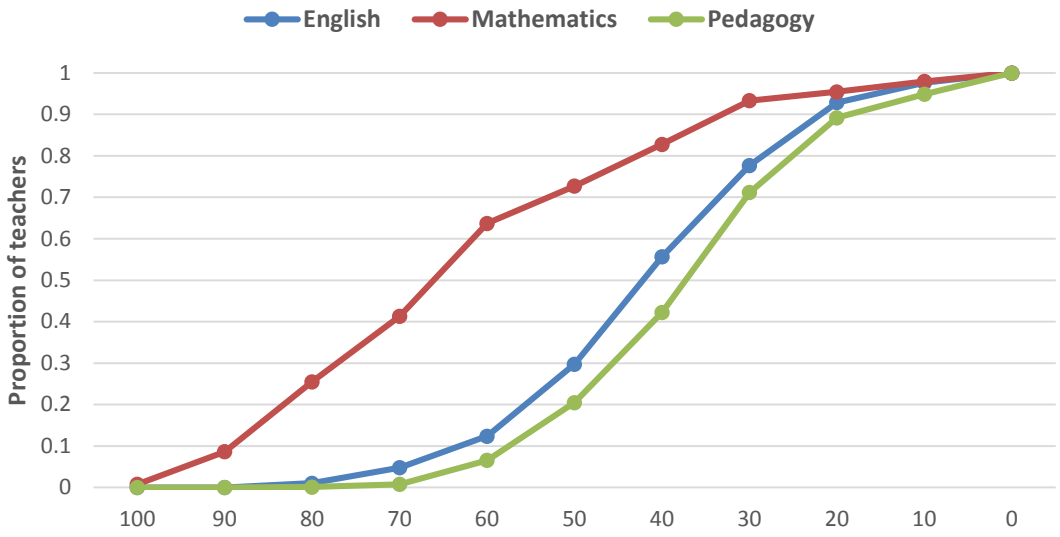
(Percent)	Tanzania	Dar es Salaam	Other Urban	Rural	EQUIP-T	All Rural	All Urban
Minimum Knowledge cut-off point: 80%	***0.1***	***0.4***	***0.0***	***0.2***	***0.1***	***0.2***	***0.1***
Average scores							
Pedagogy (complete test)	***35.9***	***42.4***	***37.1***	***35.2***	***36.5***	***35.2***	***38.0***
Basic pedagogy	32.2	38.7	32.8	31.7	31.5	31.7	33.9
Advanced pedagogy	38.7	45.2	40.3	37.9	40.3	37.9	41.2
Prepare a lesson plan	18.1	26.0	20.8	16.9	17.3	16.9	21.8
Compare/Assess pupils' abilities	57.8	62.9	57.1	57.6	60.1	57.7	58.0
Evaluate pupils' progress	21.8	28.3	24.1	20.8	21.1	20.8	25.0

Source: Tanzania SDI 2014 based on author's calculations

Sensitivity of Minimum Knowledge to the cut-off point

Figure 5 shows the sensitivity of the *Minimum knowledge* indicator to different cut-offs (i.e., requiring a score of 100 percent, 90 percent, etc.). First, note that teachers performed better in mathematics compared to English or pedagogy, no matter the standard used. The English graph also (first-order) stochastically dominates that of pedagogy, meaning that teachers fared better in English than pedagogy, no matter the standard for minimum knowledge. Overall, the results appear fairly sensitive to the choice of threshold, especially for English and pedagogy. For mathematics, the share of teachers who met the standard smoothly and steadily increases across the spectrum the lower the threshold. For English and pedagogy, however, almost no teacher met the standard until the threshold is set at 60 percent of mastery of the curriculum. At this point, 6.4 percent and 12.4 percent of the teachers scored enough to meet the standard in pedagogy and English, respectively. There is a very sharp increase in the proportion of teachers that met the standard when it is lowered from 60 percent to half and then 40 percent of the curriculum. There is a jump of 43 percentage points for English and 36 percentage points for pedagogy when the standard moves from 60 to 40 percent. As noted before, however, it seems reasonable -if not too low a bar- to request teachers to master at least 80 percent of the curriculum they teach for them to deserve to stand in front of a classroom.

Figure 5. Sensitivity of *Minimum knowledge* to the cut-off point



Determinants of teacher's performance

Are there key determinants or correlates of teachers' performance that one can identify? It is widely recognized now that what teachers know is the most important driver of what pupils learn (see Metzler, and Woessmann, 2012, among others). In turn, teachers' knowledge and performance can have long lasting impact on their pupils through labor market outcomes and productivity as shown by Hanushek (2011).

As expected, gender, education level, training, and teaching experience prominently figure among the usual suspects for teacher's quality determinants. There are clearly more variables that influence teacher's quality but the regression results in Table 17 focus on those few important aspects and check how strongly they are correlated with teachers' standardized scores in mathematics, English, pedagogy, and combination scores of those three.

Perhaps surprisingly, female teachers performed less well than their male colleagues across the board except in English. Controlling for a host of other factors, female teachers, on average, scored an astonishing 0.28 standard deviation (SD) in mathematics and 0.14 SD in pedagogy. Older teachers also scored less well with five additional years of age also lowering the mathematics score by 0.11 SD. Specialized teachers (i.e., teachers who teach only one subject) performed better than teachers who teach multiples subjects. For instance, mathematics- or English-only teachers scored 0.328 SD and 0.378 SD more on their respective subject compared to teachers who taught both subjects. However, mathematics-only teachers scored significantly more poorly in pedagogy than English or non-specialized teachers.

Because education level and teacher training are highly correlated, only the first is included in the regression analysis of Table 17. The large majority (85 percent) of Tanzania's primary school

teachers reported to have completed lower secondary school and have, therefore, finished the O-level.[17] Few but a non-negligible share (four percent) have not even transitioned out of primary level.

Table 17. Regression results of teachers' standardized test scores

	Mathematics	English	Pedagogy	Mathematics & English	Mathematics, English, & Pedagogy
Age	-0.0220***	-0.00694	-0.0202***	-0.0161***	-0.0229***
	(0.00716)	(0.00656)	(0.00523)	(0.00559)	(0.00578)
Female	-0.286***	-0.0312	-0.145***	-0.212***	-0.247***
	(0.0547)	(0.0661)	(0.0463)	(0.0392)	(0.0433)
Reference group is teaches mathematics and English					
English teacher		0.378***	0.140*	-0.316***	-0.362***
		(0.0505)	(0.0759)	(0.0470)	(0.0624)
Mathematics teacher	0.328***		-0.141**	1.007***	0.530***
	(0.0680)		(0.0606)	(0.0629)	(0.0698)
Reference group is O-Level education					
A-Level education	0.0587	0.331***	0.363***	0.170**	0.332***
	(0.0870)	(0.0899)	(0.0928)	(0.0715)	(0.0873)
Post Diploma	0.0516	0.116	0.236	0.0477	0.153**
	(0.187)	(0.129)	(0.150)	(0.0945)	(0.0685)
University	-0.0540	0.405**	0.565***	0.173	0.416***
	(0.193)	(0.177)	(0.140)	(0.116)	(0.135)
Primary	-0.398***	-0.842***	-0.320***	-0.450***	-0.534***
	(0.119)	(0.229)	(0.0934)	(0.118)	(0.107)
Year began teaching	-0.0229***	-0.0162**	-0.00682	-0.0196***	-0.0188***
	(0.00655)	(0.00773)	(0.00570)	(0.00494)	(0.00549)
Year finished teacher training	0.00715*	0.0138***	0.000586	0.00921**	0.00794**
	(0.00414)	(0.00482)	(0.00377)	(0.00350)	(0.00372)
Taught in other school last year	-0.0948	-0.121	-0.0634	-0.102*	-0.120**
	(0.0713)	(0.0811)	(0.0543)	(0.0523)	(0.0553)
Reference group is Dar es Salaam					
EQUIP-T	-0.385***	-0.362***	-0.330***	-0.379***	-0.487***
	(0.107)	(0.107)	(0.0928)	(0.0717)	(0.0951)
Other urban	0.0170	-0.290***	-0.241***	-0.138***	-0.238***
	(0.0865)	(0.0668)	(0.0903)	(0.0485)	(0.0618)
Rural	-0.234**	-0.334***	-0.397***	-0.293***	-0.448***
	(0.0991)	(0.0968)	(0.0851)	(0.0659)	(0.0715)
Constant	32.94**	5.612	13.75	21.72*	23.24*
	(15.57)	(13.13)	(10.92)	(12.39)	(12.64)
Observations	1,466	1,186	2,125	2,125	2,125
R-squared	0.078	0.132	0.084	0.387	0.211

Note: Robust standard errors in parentheses. Standard errors clustered at the school level. Levels of significance: *** p<0.01, ** p<0.05, * p<0.1.

[17] According to MoEVT's policy (MoEVT, 2007), completing A-level secondary is necessary to enroll for a teacher education diploma course. The standard for hiring teachers have clearly slipped.

Level of education has a strong impact on teachers' test performance. Teachers who reported only completing primary scored significantly worse across the board. Unlike English, mathematics scores are not very responsive to higher education with teachers with A-level education or even university degrees performing at the same level than O-level teachers. This might be due to the level of the test itself and the fact that English is not used very widely in primary schools. Finally strata dummies confirm that EQUIP-T and rural teachers are the worst performers. Dar es Salaam teachers outperform teachers in other urban areas in all subjects except mathematics.

Delays in salary and unpaid claims

It is expected that the level of effort teachers exert both by showing up at school and also maintaining a "normal" level of teaching time by spending more productive time in the classroom is highly correlated with teachers' incentives. In a system that provides proper incentives one can reasonably expect that teachers will exert higher levels of effort than shown in Table 8. In Tanzania, the salience of teachers' incentives has been fully recognized. The Big Results Now initiative (2013) report has identified the key challenge as "the teaching profession is not respected, and teachers lack motivation, accountability and commitment to deliver". The BRN Education Lab had set itself to kick start a "teacher perception transformation" through both monetary and non-monetary incentives. One important issue the BRN Education Lab set as a priority was the resolution of teachers' unpaid claims. The lab suggested that all outstanding claims should be cleared and all upcoming claims handled in a swift manner by inter alia allocating a ring-fenced budget for teachers' allowances. In addition to allowances, the most basic incentive for teachers is their salary. Delays in the receipt of salary can discourage and demotivate teachers.

Table 18 shows teachers' answers to the questions (1) "how many times have you had a delay in salary over the past 12 months?" and (2) Do you have any other unpaid claim?" Almost one out of three teacher (32.9 percent) claimed to have experienced at least one salary delay over the year preceding the survey. Teachers in urban areas other than Dar es Salaam were more likely to report such a delay. Table 18 also suggests that unpaid claims were yet to be cleared with almost half (46.7 percent) of the teachers reporting at least one unpaid claim. Dar es Salaam teachers were, again, more likely to report an unpaid claim with 70.4 percent of them doing so. The most prevalent unpaid claim nationally was salary arrears (23.7 percent) followed by leave (9.4 percent) and relocation allowance and hardship pay with 3.5 percent and 3.4 percent, respectively.

Table 18. Share of teachers' unpaid/outstanding claim (Major claims only)

(Percent)	Tanzania	Dar es Salaam	Other Urban	Rural	EQUIP-T	All Rural	All Urban
Salary delay: At least once	***32.9***	***32.2***	***40.0***	***31.0***	***29.6***	***31.0***	***38.2***
No claim	***53.3***	***29.6***	***46.8***	***57.4***	***53.6***	***57.1***	***43.0***
Unpaid leave	9.4	35.5	12.3	6.1	7.6	6.2	17.9
Hardship pay	3.4	1.6	3.3	3.5	2.7	3.5	3.0
Relocation allowance	3.5	1.5	3.8	3.6	5.8	3.6	3.2
Salary Arrears	23.7	22.7	28.4	22.4	19.3	22.6	26.6
Appointment allowance	2.0	1.2	2.4	2.0	2.9	2.0	2.1
Other claim	3.1	5.5	2.3	3.1	7.6	3.0	3.1
No answer	1.7	2.4	0.7	1.9	0.6	1.9	1.1

Source: Tanzania SDI 2014 based on author's calculations

IV. ASSESSMENT OF PUPIL LEARNING

Given that Kiswahili is the main medium of instruction in Tanzania's primary schools, a set of pupils were administered exactly the same English test except that it was translated into Kiswahili. Altogether, the language test included six tasks. All the pupils sat for the same mathematics test irrespective of the language test they took. The overall results for the English/Kiswahili and mathematics scores are reported in Table 19.

Unsurprisingly, pupils scored much higher on the Kiswahili test compared to English. Indeed, the pupils who sat for Kiswahili correctly answered 64.5 percent of questions on the overall test compared to 49.5 percent for those tested in English.[18] The average score in mathematics was 58.3 percent and is exactly the same for both groups of pupils. Therefore, language score differences fully account for the difference in average test scores. As a matter of fact, the average score in English was only 37 percent, whereas the pupils scored an average of 81 percent in Kiswahili, a very commendable performance. The pupils were also tested on four non-verbal reasoning questions and received an average score of 53.7 percent on that part of the assessment.[19]

Unlike the other indicators, a sharp pattern appears for regional inequality in the pupils' test scores. Dar es Salaam pupils significantly outperformed pupils in the other urban areas, who in turn outperformed pupils in rural areas and EQUIP-T regions. The scores for the last two were not statistically distinguishable. This pattern holds for all sections of the test. As a consequence, urban pupils scored significantly higher than rural pupils. However, girls and boys performed at par. Quite surprisingly, whether the pupil had breakfast – at home or at school - did not make a difference when compared to those who were not fortunate enough to have breakfast.

While the average total score provides interesting information, looking into the details of the test can give even more insights. Table C 4 and Table C 5 break down the results for English/Kiswahili and mathematics, respectively.

For Kiswahili, close to nine out of 10 pupils manage the simplest tasks, such as identifying a letter or recognizing a simple word. Actually, Kiswahili test takers performed well even with more complex tasks, such as reading a five-word sentence (83.1 percent) or a 50-word paragraph (74.9 percent), even though one out of four pupils could not read the paragraph. The performance significantly drops down to 63.5 percent when pupils were asked factual comprehension questions about the paragraph they read, and only 42.8 percent of them could answer a question requiring higher analytical skills.

The picture is completely different for English, with pupils unable to identify a simple alphabet letter one out four times they were asked to do so. When it comes to more complex tasks, only one out of three pupils (31.4 percent) could read a 10-word sentence and a dismal 2.7 percent could fluently read the 58-word paragraph with which they were presented.[20] Given this, it is not surprising that only around one out of eight (13 percent) of the standard four pupils could answer a factual question about the text and one out of seven (17.6 percent) could answer a question about the meaning of the passage.

[18] The total score is a simple average of all questions in the Kiswahili/English and mathematics sections. A correct answer collects one point and a wrong answer is worth zero.

[19] Just as for mathematics, the pupils performed similarly on NVR irrespective of the language test they took.

[20] Only 37 percent of the pupils could read at least half of the words in that paragraph.

Box 4. Background on the SDI Pupil Assessment

It is instructive to think of the Service Delivery Indicators as measuring key inputs, with a focus on what teachers do and know, in an education production function. These inputs are actionable and they are collected using objective and observational methods at the school level. The outcome in such an education production function is pupil learning achievement. While learning outcomes capture both school-specific inputs (e.g., the quality and effort exerted by the teachers) and various child-specific factors (e.g., innate ability) and household-specific factors (e.g., the demand for education), and thus provide, at best, reduced form evidence on service provision, it is a still an important measure to identify gaps and to track progress in the sector. Moreover, while the Service Delivery Indicators measure inputs -- and learning outcomes are not part of the Indicators -- in the final instance we should be interested in inputs not in and of themselves, but only in as far as they deliver the outcomes we care about. Therefore, as part of the collection of the Service Delivery Indicators in each country, learning outcomes are measured for grade four pupils.

The objective of the pupil assessment was to measure basic reading, writing, and arithmetic skills. The test was designed by experts in international pedagogy and based on a review of primary curriculum materials from 13 African countries (For details on the design of the test, see Johnson, Cunningham and Dowling (2012) "Draft Final Report, Teaching Standards and Curriculum Review"). The pupil assessment also measured nonverbal reasoning skills on the basis of Raven's matrices, a standard IQ measure that is designed to be valid across different cultures. This measure complements the pupil test scores in language and mathematics and can be used as a rough measure to control for innate pupil ability when comparing outcomes across different schools. Thus, the pupil assessment consisted of three parts: language, mathematics and non-verbal reasoning (NVR).

The test, using material up to the grade three level was administered to grade four pupils. The reason for choosing pupils in grade four is threefold. First, there is scant information on achievement in lower grades. SACMEQ, for example, tests pupils in grade six. Uwezo is a recent initiative that aims to provide information on pupils' learning irrespective of whether they are enrolled in school or not and tests all children under the age of 16 on grade two material. While this initiative has provided very interesting results, it is not possible to link pupil achievement to school level data, since the survey is done at the household level. Second, the sample of children in school becomes more and more self-selective as one goes higher up due to high drop-out rates. Finally, there is growing evidence that cognitive ability is most malleable at younger ages. It is therefore especially important to get a snapshot of pupil learning and the quality of teaching provided at younger ages.

The test was designed as a one-on-one test with enumerators reading out instructions to pupils in their mother tongue. This was done to build up a differentiated picture of pupils' cognitive skills; i.e. oral one-to-one testing allows us to evaluate whether a child can solve a mathematics problem even when her reading ability is so low that she would not be able to attempt the problem independently. The language test consisted of a number of different tasks ranging from testing knowledge of the alphabet, to word recognition, to a more challenging reading comprehension test. Altogether, the test included six tasks. The mathematics test also consisted of a number of different tasks ranging from identifying and sequencing numbers, to addition of one- to three-digit numbers, to one- and two-digit subtraction, to single digit multiplication and divisions. The mathematics test included six tasks and a total of 17 questions. The non-verbal reasoning section consisted of four questions.

Table 19. Tanzania standard four pupil performance

(Average scores in percent)	English & mathematics	Kiswahili & mathematics	English	Kiswahili	Mathematics	NVR
Tanzania	40.1	76.2	36.5	80.9	58.2	53.6
Dar es Salaam	65.0	92.6	63.2	97.5	75.0	67.7
Other Urban	52.1	84.8	50.3	90.4	62.1	57.4
Rural	35.2	72.8	31.1	77.3	55.9	51.5
EQUIP-T	32.6	69.3	29.0	74.0	50.6	47.3
All Rural	35.5	73.0	31.4	77.5	56.1	51.8
All Urban	55.0	86.5	53.3	91.9	65.0	59.2
Boy	41.7	77.8	38.0	82.5	60.1	55.3
Girl	39.1	74.7	35.6	79.4	56.9	52.4
Had Breakfast	40.0	77.9	36.5	83.1	58.2	52.3
No Breakfast	40.5	75.1	36.9	79.5	58.5	54.6
Number of pupils	2,813	1,188	2,813	1,188	3,983	3,983

Source: Tanzania SDI 2014 and author's calculations

Mathematics scores were slightly better than the English scores, but below the Kiswahili performance. There were still some significant knowledge gaps. In terms of operations, the pupils performed better when handling tasks involving only one-digit numbers, except for addition where 60.2 percent of the pupils could add two three-digit numbers. In order of performance, mastery of addition was followed by subtraction, then division as a distant third and, finally, multiplication closely followed. For operations involving two-digit numbers, 39.3 percent of the pupils could do subtraction but this performance dropped down to 18.1 percent for division, and 12.1 percent for multiplication. Very few pupils could perform on questions that required higher analytical skills, such as completing a sequencing of numbers with a specific pattern (14.1 percent) or problem-solving task (8.9 percent).

Although the pupils correctly answered more than half of the mathematics questions, the test revealed that the majority of standard four pupils did not perform well at the standard three level. For example, the complete 9x9 multiplication table is intended to be taught by standard three; simple division is also clearly in the curriculum. However, only 40 percent of the standard four pupils could perform 6÷3 or 7x8.

Pupils in Dar es Salaam schools performed significantly better across the board. They were followed by pupils in other urban areas, who also performed significantly better than rural and EQUIP-T pupils. Dar es Salaam pupils correctly answered 34 percent and 25 percent more questions than EQUIP-T pupils in English and mathematics, respectively. They also outperformed pupils in rural schools on non-verbal reasoning by 16 percentage points. For division involving two-digit numbers, Dar's pupils outperformed EQUIP-T pupils by almost four to one. For more complex operations, such as problem-solving or completing a sequence, all pupils performed poorly. But pupils in Dar es Salaam schools were five times more likely to get it right than EQUIP-T school pupils (23.8 percent vs. 4.2 percent and 26.9 percent vs. 6.7 percent, respectively).

Interestingly, socio-economic status and gender seemed to have no impact on pupils' performance. Indeed, pupils who had breakfast before coming to school, who, on average, probably live in better-off households, performed at par with children who showed up at school with an empty stomach. There was also no noticeable difference between boys and girls.

Figure 6. Performance correlation reading a paragraph and mathematics by language

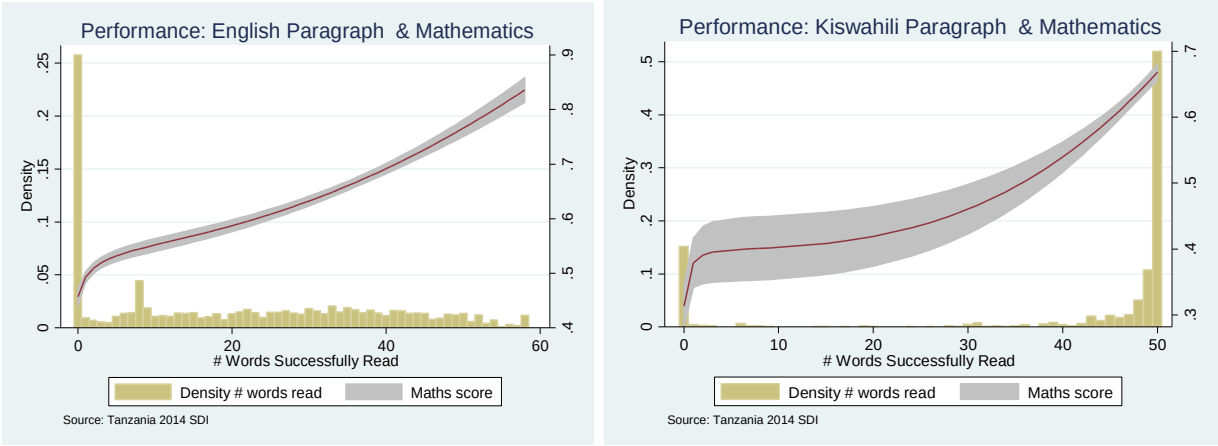


Figure 6 shows how performance in reading a language paragraph relates to performance in mathematics. The histograms show the distribution of pupils by the number of words they could read in the paragraph (58 words for English and 50 for Kiswahili). The lines show the average mathematics score (on the right-hand side, vertical axis) against the number of words successfully read by the pupils. The distribution of number of words read in Kiswahili shows some pupils bunching on zero, then very few pupils reading between one and 40 words, and, finally, the majority of pupils reading on the higher end of the distribution and the biggest bunching on a perfect score (actually, 53 percent of children correctly read the entire Kiswahili paragraph). For English, the bunching is on zero with around 26 percent of the pupils who could not read a single word; then, the distribution is roughly uniform across the remaining of the support, meaning a more or less similar number of kids could read 10 or 17 or 35 words in the paragraph.

The graphs clearly show that mathematics and language are strongly positively correlated. This means the more words the pupils read correctly, whether English or Kiswahili, the better they performed in mathematics. The grey area depicts the 95 percent confidence interval around that score. The mathematics scores are fairly precisely estimated across the board for English test takers, whereas the confidence intervals are large for Kiswahili. This happens mostly in the middle range because pupils scored either very poorly or reasonably well on Kiswahili and are, therefore, at the extreme of the distribution. There are also more children tested in English than Kiswahili. The graphs also show that, even though Kiswahili and English test takers performed similarly on average, the best performers in English (i.e., those who were able to read all or almost all the words) performed better in mathematics than their peers in Kiswahili.

Correlations between service delivery indicators and learning outcomes

As stated earlier, the service delivery indicators are, by no means, an end in and of themselves. In fact, they matter if, and only if, they explain and are somehow related to pupils' learning, which is what the population and the policy-makers care most about. Table C 6 to Table C 10 show a series of simple regressions of pupils' test scores for various subjects on individual service delivery indicators. The scores, as well as the indicators, are averaged at the school level and standardized.

The top panel (Panel A) shows the regression results for all the schools, whereas the sample is split into urban and rural schools for the middle panel (Panel B) and bottom panel (Panel C), respectively. All the indicators are significant in one regression or another and they always have the expected sign. The coefficients also suggest that the indicators have relatively large effects; for example, pupils' mathematics scores increased by 19 percent (resp. 15 percent) of a standard deviation, if teachers' scores in mathematics increased by one standard deviation or if class size was reduced by one standard deviation (see Table C 8). Infrastructure and time spent teaching also come out with strong effects on learning.

V. SDI SPECIAL TOPIC: GENDER

Gender equality is a leading Millennium Development Goal (MDG 3) and its first sub-component is about equality in education. The second MDG also is about gender equality because it advocates for universal primary education. There is evidence that gender equality is a strong driver for economic growth and, as put in simple terms by the 2012 WDR, gender equality is smart economics. Gender equality in education is most probably the strongest route to full gender equality. This goal has been embraced by most countries and many have reached the MDG, including Tanzania. Unfortunately, though, equality in education is generally equated to having equal numbers of boys and girls in school. Although ensuring that girls have as much access to education as boys is a worthy goal, it may not be enough. One needs also to make sure that girls are treated in the same way as boys in schools and classrooms and that they have equal access to inputs, be it books, stationery, toilets, or teachers' time and attention.

To the best of our knowledge, this is the first attempt in SDI and other facility surveys in Africa to produce gender-disaggregated school and classroom observation data.[21] The SDI survey collected the complete teacher roster in each and every visited school. Questions were asked of all teachers, but subsequently a number of teachers have been sampled for investigating absence rates, knowledge, and classroom practices. In its classroom observation module, the 2014 Tanzania Education SDI survey collected detailed gender-disaggregated data. This section uses the collected information to analyze and shed new light on gender issues in Tanzanian primary schools and classrooms.

A. Teachers' characteristics by gender of head teacher

Tanzania has succeeded in achieving gender parity, not only for pupils, but also for teachers in its primary education system. Table 20 indeed shows that women constituted half (49.8 percent) of primary school teachers body. There was quite a lot of regional variation hidden in this average. For instance, in Dar es Salaam an overwhelming majority (84 percent) of teachers were female. The same was true in other urban areas, albeit with a smaller but still strong majority, with 74 percent of female teachers. In contrast, rural schools were still lagging behind in teacher gender parity with only two-in-five female teachers (40 percent).

Table 20. Share of female teachers in school

	By Position		By Sex of Head Teacher	
(Percent)	All Teachers	Head teachers	Male	Female
Dar es Salaam	84.4	46.6	81.2	87.9
Other Urban	74.0	37.0	66.3	78.9
Rural	40.0	14.5	36.7	57.7
Total	49.8	17.9	43.7	69.6

[21] The SACMEQ data actually includes a variable on using a book in the classroom, but this is self-reported as it is asked directly to the pupil instead of direct observation as is the case with SDI.

Source: Author's calculations using 2014 Tanzania SDI data

In school management, however, females were grossly under-represented with less than one out of five (18 percent) of the head teachers a woman. Even though they had the same weight (in terms of number of people) in the teaching force, women are five times less likely to head a school compared to men. Rural areas displayed the biggest deficit of female school managers (14 percent), but it was in Dar es Salaam that the gender gap (35 percent) was the most important.[22]

Female-headed schools had on average 25.9 percent more female teachers a strong and significant difference. As Table 20 shows, on average, seven out of 10 teachers in schools managed by a woman were female compared to slightly more than four out of 10 (44 percent) in schools headed by men. Women, in fact, constituted the majority of the teaching force in all schools, except rural ones that are headed by men. In the latter schools, there were a 2:1 male to female teacher ratio. It is also interesting to note that in terms of teacher staff size, female-headed schools were significantly larger than those headed by men. This held true across all strata (Figure 7) and the average female head of school had to manage 3.9 more teachers than her male counterpart.

Figure 7. Female-headed schools have more teachers

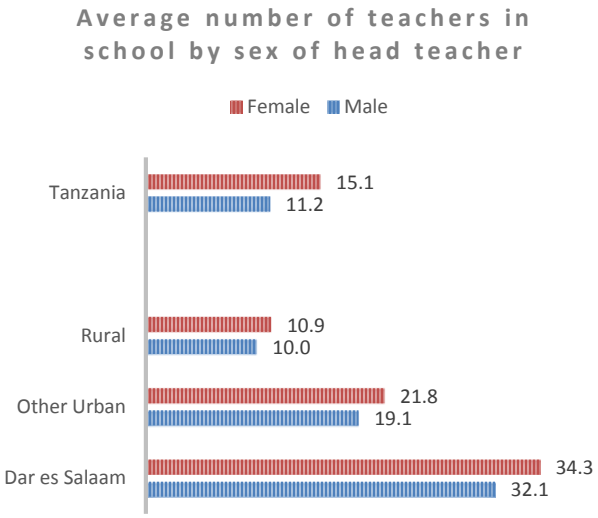


Table 21. Teachers' completed level of education

	Male head			Female head		
	M	F	All	M	F	A
Primary	4.6	3.9	4.3	5.7	2.2	3.2
O-Level	85.4	88.4	86.7	81.2	84.0	83.2
A-Level	6.1	2.9	4.7	6.8	5.0	5.5
Post-Secondary	3.7	**4.1**	3.8	6.2	**8.6**	7.9

	All head teachers		
	M	F	All
Primary	4.7	3.3	4.1
O-Level	84.8	87.0	85.9
A-Level	6.2	3.6	4.9
Post-Secondary	4.0	5.5	4.8

Source: Author's calculations using 2014 Tanzania SDI data

Overall, the education profile of male and female primary school teachers was fairly similar, as shown in Table 21. The vast majority of primary school teachers (86.7 percent) have just completed the ordinary secondary level (i.e., four years of post-primary). The remaining 14 percent of teachers were almost equally distributed across the other education levels with four percent who completed primary (i.e., the level they teach), five percent completed the advanced secondary level, and four percent went beyond secondary level.

The distribution of teachers across schools seems, however, to depend on the sex of the head teacher. Indeed, it appears as if teachers who have gone beyond the O-level were more likely to end up in a female-headed school. This is even stronger for female teachers, 9.1 percent of whom have gone beyond O-level but 13.6 percent were in female-headed schools vs. seven percent in male-

[22] The gender gap is measured as the difference between proportion of female in the teaching force and the share of female in school managerial positions.

headed school. Better educated female teachers, who are likely to have more say in their options, seemed to choose to teach in female-headed schools.

Does the sex of the head teacher also influence teachers' behavior? We already learned in an earlier section of this report that male and female teachers were equally likely to be absent (from school and the classroom). This is confirmed in Table 22, but it also shows that male teachers behaved differently in school depending on the sex of the head teacher. Male teachers were almost 10 percentage points more likely to be in the classroom teaching when the head teacher was a man compared to a woman. This difference is entirely explained by the fact that the teachers were in school, but not in the classroom. Female teachers, in contrast, displayed, by and large, the same behavior irrespective of the sex of the head teacher. Table C 11 offers more results on teachers' practices in the classroom broken down by gender.

Table 22. Male teachers more likely in classroom teaching when head teacher is male

	Male head teacher			*Female head teacher*			*All head teacher*		
	Male	Female	Total	Male	Female	Total	Male	Female	Total
In class - teaching	**50.1**	48.4	49.4	**41.5**	49.6	47.2	48.9	48.8	48.9
In class - not teaching	4.4	3.9	4.2	3.7	4.4	4.2	4.3	4.1	4.2
In school - not in class	**31.9**	31.5	31.7	**40.1**	30.8	33.6	33.0	31.2	32.1
Absent from school	13.6	16.2	14.7	14.7	15.2	15.0	13.7	15.9	14.8

Source: Author's calculations using 2014 Tanzania SDI data.

Finally, we looked into the intensity of classroom supervision (Figure 8) and the incidence of salary delays and unpaid claims (Table 23) by the sex of teachers and head teachers. It is noteworthy that most teachers (54 percent) said that the head teacher comes to observe their classroom at least once a week. Male head teachers seemed to scrutinize their female teachers more intensively. In contrast, there is no noticeable difference in classroom observation by female head teachers, who treated all teachers equally in terms of classroom observation.

Figure 8. Supervision of head teacher

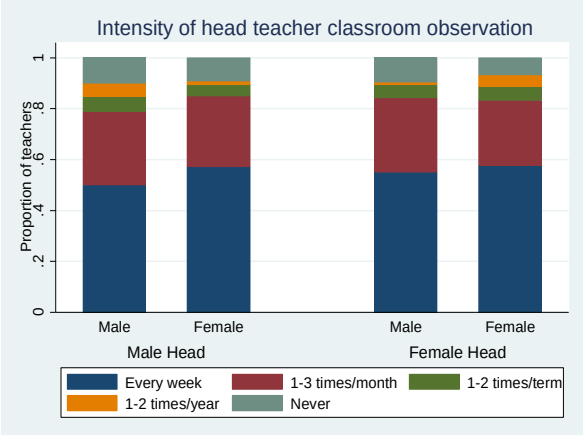


Table 23. Unpaid claims and salary delays

	Reported Unpaid Claims (%)			
	Male	Female	Diff.	P-value
All head	45.4	51.5	6.1***	0.0001
Male head	45.5	50.0	4.5***	0.0077
Female head	45.2	54.6	9.4***	0.0066
	Reported Salary Delays (%)			
All head	32.3	33.5	1.2	0.2161
Male head	32.3	33.4	1.1	0.2646
Female head	32.3	33.7	1.4	0.341

Source: Author's using Tanzania SDI 2014 data. Levels of significance: *** p<0.01, ** p<0.05, * p<0.1.

Many teachers reported receiving their salary with delay (roughly one in three) and unpaid claims (roughly one in two). There is, however, no significant difference in salary delays with male and female teachers equally likely to have experienced a delay and irrespective of the sex of their head teacher. However, there were significant difference in terms of unpaid claims. About 45 percent of male teachers reported an unpaid claim. This proportion is the same whether the head teacher was a man or a woman. Female teachers, on the other hand, were significantly more likely (6.1 percent) to report an unpaid claim. In addition, female teachers in a female-headed school were the ones most likely to report unpaid claims with a 9.4 percent significant difference with male teachers working under a female head.

B. Gender effects of classroom environment and teaching practices

In each school, one standard four classroom was observed during an English or mathematics lesson. The objective of the classroom observation was to have a better understanding of the dynamics in a typical standard four class. Information was collected on the classroom environment and how the teacher carried out her teaching activity (i.e., how she behaved with pupils, whether she asked questions, provided feedback, went individually to the pupils, called them by name, etc.). Most of the questions were gender-disaggregated (for instance, the enumerator was asked to note the pupil's gender when recording the number of pupils who went to the blackboard).

Table 24. Class size and use of inputs by pupils in classroom

	Male Teachers				Female Teachers			
	Mean	Std. Err.	[95% Conf. Interval]		Mean	Std. Err.	[95% Conf. Interval]	
Total number of pupils	**39.7**	**1.7**	**36.3**	**43.1**	**48.6**	**2.6**	**43.4**	**53.8**
Of which % girls	53.5	0.9	51.6	55.4	53.7	0.9	52.0	55.4
Share of pupils with textbook	**22.9**	**2.7**	**17.5**	**28.2**	**23.3**	**3.3**	**16.9**	**29.7**
Of which % girls	47.9	2.4	43.3	52.6	54.6	2.3	50.0	59.2
Share of pupils that used textbook	**38.6**	**4.0**	**30.7**	**46.5**	**37.7**	**4.2**	**29.5**	**45.9**
Of which % girls	48.8	2.2	44.5	53.1	52.7	2.0	48.7	56.6
Share of pupils with pens	**95.9**	**0.9**	**94.1**	**97.7**	**92.6**	**1.8**	**89.0**	**96.2**
Of which % girls	52.8	1.0	50.7	54.8	52.7	1.0	50.7	54.7
Share of pupils with exercise books	**95.9**	**1.2**	**93.6**	**98.3**	**95.6**	**0.9**	**93.9**	**97.3**
Of which % girls	52.8	1.0	50.8	54.8	52.7	1.0	50.8	54.6

Source: Author's calculations using 2014 Tanzania SDI data

Teachers led and maintained gender-balanced classrooms in terms of access to and use of inputs (Table 24). Female teachers had larger classes with a difference of almost nine pupils. The gender composition of the classes was, however, similar with roughly 54 percent of girls in the classroom irrespective of the teacher's gender. In terms of access to inputs such as textbooks, pens or pencils, or exercise books, girls and boys had access in the same proportion as they were in the class (i.e., of the pupils who had access to the inputs roughly 52 to 54 percent are girls). The major exception concerned textbooks, with girls being slightly at a disadvantage with male teachers. To illustrate, in the average male teacher's classroom, only 23 percent of the pupils had a textbook and among those, 48 percent were girls. However, girls represent 54 percent of the classroom population.

This difference was not apparent when the teacher was female. There was no difference either for pens or exercise books, but this could be a reflection of the near universal access to those inputs with 96 percent of pupils having them. As they were in short supply, books were widely shared among pupils. Indeed, if only 25 percent of the pupils could show they had the textbook handy, 39 percent used it while in the classroom.

Table 25 shows regression results for many teaching practices as observed in the classroom. (For a fuller breakdown of teaching practices, see Table C 12.) Each row represents a regression. The first three columns show combinations of head teacher/teacher by sex. The reference group is HM-TM, whereby both the head teacher and the teacher are male. The fourth column is the share of female teachers in the school. Finally, the regressions control for stratum with Dar es Salaam as the reference region. By and large, the teachers' interaction with their classroom seemed quite gender-balanced.

Table 25. Regression results of teaching practices

	HF-TM	HM-TF	HF-TF	% female teachers	Other urban	Rural	# Obs.	R-squared
Number pupils in classroom	-2.097	**2.513***	**-4.099***	16.58	-4.236**	-22.44**	396	0.151
% of girls in classroom	-0.0251	0.00840	0.00866	0.0174	0.00937**	0.0303**	396	0.013
Share of pupils with textbooks	**0.138***	**-0.037****	0.0221	**0.159*****	-0.129***	0.0183	395	0.042
% of girls with textbooks	0.0934	**0.0573****	**0.0870****	**0.121****	0.0320***	0.0852***	233	0.063
Share of pupils who used textbooks	**0.0762*****	**-0.019****	**0.131***	0.0156	-0.149***	0.000914	392	0.029
% of girls using textbook	0.0234	0.00927	0.0169	0.108	0.0348*	0.0212	214	0.051
Teacher wrote on board	0.0111	0.00392	-0.00286	-0.0104	-0.016***	-0.015***	394	0.003
Pupils wrote on board	-0.115	**0.106****	0.145	-0.317	-0.00345	-0.00916	394	0.022
% of girls wrote on board	-0.0637	-0.100	**-0.088*****	**0.347****	0.135***	0.0872*	170	0.065
Teacher kept attendance	0.0259	0.0410	0.0468	0.0246	0.0149	0.0693	396	0.008
Teacher had scheme of work	0.0106	-0.0142	-0.0271	**0.199****	0.107***	0.0922**	396	0.012
Teacher had lesson plan	-0.153	-0.0707	-0.0681	**0.394****	0.00293	-0.00211	396	0.050
Teacher introduced lesson	-0.0515	0.00399	-0.0167	-0.0657	-0.0177**	-0.113***	389	0.013
Teacher summarized lesson	0.0116	0.0889	0.0267	-0.102	0.146***	0.0169	393	0.014
Teacher assigned homework	-0.150	**0.0848****	0.186	**-0.315***	-0.173***	-0.0677	393	0.038
Teacher tested creativity	**0.271****	**0.144****	0.0637	-0.0104	0.0925***	0.0667*	394	0.026
Teacher gave positive feedback	**0.131***	0.103	**0.160*****	-0.0314	-0.0503*	-0.0291	394	0.022
Teacher gave corrective feedback	0.00507	0.0321	**0.0759****	-0.0755	0.0540***	-0.0302	394	0.009

Note: Significance levels *** p<0.01, ** p<0.05, * p<0.1.

There is no discernable gender effect in keeping attendance, having a scheme of work, introducing or summarizing the lesson, or providing feedback (positive or negative) to pupils. Teachers also visited pupils or sent them to write on the blackboard with no differences in gender from the teachers' or the pupils' side. There are, however, a few interesting exceptions. The textbook-gender imbalance effect noted in Table 24 is confirmed by the regressions (rows 4 to 6). Girls are significantly more likely to have access to a textbook when they have a female teacher compared to a male teacher regardless of the sex of the head teacher although their likelihood to access a textbook is enhanced when the head of the school is a woman as well. A girl with a female teacher in a female-headed school had 8.7 percent more chance to have a book than her similar peer with a male teacher in a male-headed school. The share of female teachers is also a strong driver for a higher share of girls using textbooks or writing on the board. Finally, it is intriguing that when the head teacher and the teacher were of opposite sex, the teachers were more likely to ask questions that required the pupils to use their creativity and imagination.

C. Pupils' performance and teacher's gender

Does teacher's gender matter for pupil's learning outcomes? **Do girls perform better when they are taught by a woman?** There is a large literature on this topic (see Dickerson et al 2015, Holmund and Sund 2008, and references therein). If teachers treat pupils differently according to both the teacher's and the pupil's gender or pupils themselves behave according to their teacher's gender, this may impact pupils' learning achievement. The impact may also be cumulative along the life of the pupils through their primary cycle. A clear cut answer to this question may have deep policy implications, but our data can only offer suggestive evidence for Tanzania.

Table 26 shows the results of a regression of pupils' English or mathematics performances on a series of variables of interest. It shows the effects of having a female teacher on all pupils' test performance, as well as on boys and girls taken separately. The teachers' tests scores are also of special interest. A number of school-level variables such as the share of female teachers or whether the school is headed by a woman are also examined.

The first three columns show achievement in mathematics, whereas the last three are concerned with the English test scores. The first regression was pooled for all pupils. The sample was, then, split between boys and girls and the same model was run on each single-sex sample. The teacher is the pupil's teacher for the specific subject. Also we consider teacher's scores on both subject.

Looking into mathematics achievement, column 1 shows a gender gap of 13 percent of a standard deviation to the detriment of girls. Having a female teacher impacted negatively on pupils' mathematics scores. Although, columns 2 and 3 show that this impact was only for boys, meaning that boys performed better in mathematics when their teacher was a man. The teacher's performance on both English and mathematics had a strong positive effect on pupils' mathematics scores. Again, the single-sex regressions show that this impact appeared only for girls, meaning the better the teacher performed, the better the girls' mathematics performances, whereas this had surprisingly no bearing on boys' achievements. Finally, although the head teacher's gender had no effect on the pupils' scores, the share of women teachers in the school strongly impacted the pupils' performances, irrespective of their sex, although the effect is stronger for girls.

The results for pupils' performances in English are fairly similar to those in mathematics. However, there are a few noticeable differences. For instance, a teacher's mathematics performance had no impact on pupils' English achievement. Female teachers did not impact pupils' scores, irrespective of their gender. Finally, the more female teachers in the school, the better girls performed in English, but this had, again, no impact on boys' performances.

Table 26. Determinants of pupils' performance (standard deviations)

	Mathematics			*English*		
	All	**Boys**	**Girls**	**All**	**Boys**	**Girls**
	(1)	(2)	(3)	(4)	(5)	(6)
Pupil's characteristics						
Pupil is girl	-0.132*			-0.101*		
	(0.0434)			(0.0333)		
Pupil's age	-0.0144	0.0161	-0.0507*	-0.0302	-0.00923	-0.0546
	(0.0107)	(0.0108)	(0.0132)	(0.0204)	(0.0200)	(0.0253)
Pupil had breakfast	-0.0300	-0.0890	0.0181	-0.0876	-0.0907	-0.0802
	(0.0441)	(0.0591)	(0.0481)	(0.111)	(0.110)	(0.108)
Subject Teacher's characteristics						
Female	-0.141*	-0.180**	-0.102	0.0935*	0.0892	0.0983
	(0.0433)	(0.0283)	(0.0564)	(0.0267)	(0.0664)	(0.0415)
Age	0.000796	-0.00344	0.00379*	-0.000735	-0.00239	0.000441
	(0.000945)	(0.00128)	(0.00103)	(0.00198)	(0.00625)	(0.00273)
English score	0.0636***	0.0386	0.0853**	0.103***	0.0901**	0.109**
	(0.00633)	(0.0148)	(0.0194)	(0.00931)	(0.0116)	(0.0141)
Mathematics score	0.0536*	0.0250	0.0788**	-0.00937	-0.0531**	0.0212
	(0.0145)	(0.0117)	(0.0146)	(0.00887)	(0.00604)	(0.0170)
School characteristics						
Female head teacher	-0.0464	0.0491	-0.129	-0.0592	-0.0767	-0.0402
	(0.0639)	(0.0599)	(0.0952)	(0.0959)	(0.146)	(0.0578)
% Female teachers	0.626***	0.539**	0.718**	0.595**	0.421	0.780**
	(0.0306)	(0.0722)	(0.0779)	(0.129)	(0.188)	(0.110)
Other urban	-0.416***	-0.380***	-0.465***	-0.394***	-0.487***	-0.299***
	(0.00652)	(0.00799)	(0.00820)	(0.0181)	(0.00901)	(0.0280)
Rural	-0.522***	-0.508***	-0.538**	-0.732**	-0.799**	-0.650**
	(0.0464)	(0.0156)	(0.0680)	(0.0998)	(0.123)	(0.0768)
Constant	0.342***	0.117	0.510***	0.670**	0.577	0.677*
	(0.00572)	(0.102)	(0.0402)	(0.0962)	(0.203)	(0.220)
Observations	3,009	1,441	1,568	2,154	1,038	1,116
R-squared	0.102	0.079	0.137	0.186	0.162	0.219

Note: Robust standard errors in parentheses. Levels of significance: *** p<0.01, ** p<0.05, * p<0.1. Regressions control for other pupil's and teacher's characteristics, as well.

VI. SDI COMPARATIVE ANALYSIS

After the SDI pilot in Senegal and Tanzania was carried out in 2010, the SDI was revised and rolled out in a number of countries such as Kenya and Uganda (2013), Nigeria, Togo, and Mozambique (2014). Tanzania was, however, the first country to implement a repeat survey that would allow for trend analysis. This second SDI has a great deal of overlap with the pilot, although there were a few indicators which were not comparable. It was, however, fully comparable to the 2013 and 2014 SDIs. Teachers and pupils were assessed with the exact same questions apart from minor changes, especially for names to contextualize the survey instruments. The same questions were also asked to heads of schools. Finally, during the analysis, the indicators were computed with the same program maximizing the comparability between Tanzania and the other SDI countries.

A. Comparing Tanzania to itself: SDI trends

Tanzania, along with Senegal, was a pioneer country in the SDI by piloting the first SDI survey in 2010. The instruments have been slightly updated since then, but the 2010 and 2014 SDIs have a great deal of overlap. Some indicators were not comparable and the data necessary to compute a new one may not have been available, but, by and large, it was possible to look into the trends in service delivery between 2010 and 2014.

Table 27 shows how education indicators have evolved since the pilot SDI. In terms of access to basic infrastructure, there was no noticeable improvement in Tanzanian primary schools in the four years since the first SDI was carried out. Only two percent of the schools had electricity, clean water, and improved sanitation in 2014, down from three percent in 2010. Although urban schools seemed to have better infrastructure, they were actually statistically not distinguishable from their rural counterparts. As in 2010, the main constraint was again electricity, which was rarely available in the schools. Access to clean water, or its lack thereof, was also an issue across the board, but it was much more severe in rural schools.

Table 27. SDI Trends: How did Tanzania fare between 2010 and 2014*

	TZ SDI 2010			TZ SDI 2014		
	Tanzania	*Rural*	*Urban*	*Tanzania*	*Rural*	*Urban*
At the School						
Infrastructure*	0.03 (0.02)	0.02(0.01)	0.08 (0.08)	0.02 (0.01)	0.01 (0.01)	0.1 (0.05)
Share of pupils with textbook	0.20 (0.02)	0.17 (0.02)	0.26 (0.06)	0.25 (0.03)	0.27 (0.03)	0.18 (0.03)
Pupils in classroom	52.0 (1.9)	48.9 (2.1)	59.4 (3.9)	43.6 (1.3)	40.8 (1.6)	60.3 (3.4)
Teachers' Effort						
School absence	0.23 (0.02)	0.20 (0.02)	0.36 (0.04)	0.14 (0.01)	0.14 (0.01)	0.14 (0.01)
Classroom absence	0.53 (0.03)	0.50 (0.02)	0.68 (0.05)	0.46 (0.01)	0.47 (0.01)	0.45 (0.02)
Time spent teaching (minutes)	124 (10)	131 (10)	84 (18)	167 (9)	169 (10)	160 (15)
Teachers' Minimum Knowledge						
English & Mathematics	0.10 (0.02)	0.09 (0.02)	0.13 (0.04)	0.21 (0.01)	0.20 (0.02)	0.23 (0.02)
English	0.03 (0.01)	0.03 (0.01)	0.03 (0.01)	0.01 (0.01)	0.01 (0.01)	0.02 (0.01)
Mathematics	0.17 (0.03)	0.15 (0.03)	0.23 (0.08)	0.25 (0.02)	0.26 (0.03)	0.23 (0.03)

Note: (*) For infrastructure, we used the 2010 SDI definition (i.e., school has access to electricity, water, and sanitation). This was because, in 2010, the survey did not have lux meters to measure luminosity in the classrooms. For 2010, the minimum knowledge indicators were computed using the 2014 definition on the 2010 test data, which explains differences with the 2010 SDI report. Source: Tanzania SDI 2010 and 2014 based on author's calculations.

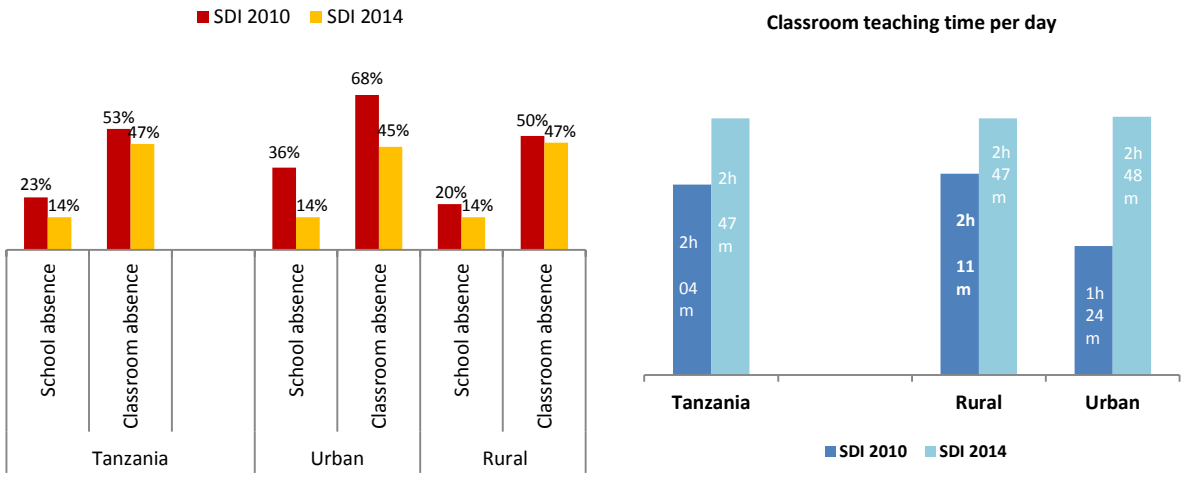
Class sizes, at least for the observed standard four classroom, dropped significantly from 52 pupils in a classroom down to 43. This amounts to a reduction of almost 20 percent, despite a growing population and a growing demand for education. This improvement is almost entirely accounted for by the reduction in the size of rural classes. As a cautionary note, it is important to remember that the indicator captures the number of pupils in the classroom (i.e., attendance) which may be lower than the number of enrolled pupils.

The most impressive progress registered in Tanzania is with teachers' effort, especially school absence which dropped from 23 percent in 2010 to 14 percent in 2014, a 40 percent decrease.

Figure 9 shows that urban schools deserved all the credit for this achievement. It is, however, disappointing to note that improvement in school absence did not fully translate into classroom absence. As a matter of fact, classroom absence decreased from 53 percent to 46 percent, but the decline was not statistically significant and about half of the teachers were still not in the classroom. The decline in urban areas was, however, significant and large, entirely due to the decline in school absence. These findings are interesting because they show that it is very important to look into what happens inside the schools. Indeed, among the teachers who showed up at school, the percentage of those who were in classroom teaching has barely change between 2010 and 2014. Strengthening the head teacher, in terms of leadership and management skills, seems of crucial importance for the better working of Tanzanian primary schools.

The reduction in school and classroom absence benefitted the pupils, who between 2010 and 2014, gained 43 minutes of teaching time per day. Over the school year, which is on average 194 days, this was a gain of 24 days of full teaching, or more than a month. Urban pupils received an impressive gain of teaching time; back in 2010, they only had 47 days of effective contact time with their teachers and this doubled to 94 days in 2014. Despite this impressive improvement, one must still keep in mind that Tanzanian primary pupils are still losing more than half of the teaching time owed to them by the education system.

Figure 9. Trend in school absence, classroom absence, and time spent teaching



Teachers' subject content knowledge in 2014 was low and has not changed compared to its 2010 level. Table 28 and Table 29 show trends in standard four teachers' and pupils' performance in test scores, respectively.[23] Teachers' performance was lowest in English and seems to have worsened between 2010 and 2014. It was mainly teachers in rural school who experienced a decline in their English skills. In mathematics, teachers scored on average 63 percent in 2010 compared to 65 percent in 2014, which is not a significant difference. Teachers scored well in simple arithmetic, such as two-digit addition, and less so in more complex tasks, like operations with fractions. There is no significant difference between 2010 and 2014 on the specific themes or topics, except for operations involving monetary units where teachers' scores have significantly declined.

Table 28. SDI 2010-2014 Tanzania Trends: Teachers' knowledge

	TZ SDI 2010			TZ SDI 2014		
(Percent)	*Tanzania*	*Rural*	*Urban*	*Tanzania*	*Rural*	*Urban*
***English** (average score)*	*57*	*57*	*57*	*42*	*41*	*46*
Grammar task	73	74	73	*73*	*72*	*78*
Reading comprehension	24	24	22	*21*	*21*	*24*
***Mathematics** (average score)*	*63*	*63*	*64*	*63*	*63*	*65*
Adding two double digits	97	98	93	*97*	*97*	*98*
Subtracting two double digits	90	90	90	*86*	*86*	*86*
Multiplying two double digits	88	89	85	*85*	*85*	*85*
Division task	88	89	85	*81*	*81*	*77*
Fractions	66	65	70	*58*	*58*	*59*
Square root (no remainder)	85	85	86	*79*	*79*	*81*
Monetary units task	70***	69***	77***	*55*	*55*	*56*
One-variable algebra	56	55	58	*50*	*50*	*51*

Note: Levels of significance: *** p<0.01, ** p<0.05, * p<0.1. Source: Tanzania SDI 2010 and 2014 based on author's calculations.

Unlike their teachers, standard four pupils' test performance showed improvement in some areas. Pupils' average English scores did not significantly change, but they performed better in reading a paragraph and in reading comprehension. Still, pupils' English performance was not good with an average score around 40 percent. Pupils' overall scores in mathematics improved significantly. When compared to their older 2010 cohort, standard four pupils in 2014 performed similarly for simple tasks, such as adding double digits or subtracting single digits, but outperformed the 2010 cohort in more complex operations, such as multiplying two triple digits or dividing a double-digit number by a single digit. Pupils in urban schools had the most significant improvement.

[23] In terms of content, the tests did not change much between 2010 and 2014. However, the teachers' test was improved in its presentation. In 2010, it looked just like a test whereas, in 2014, the teacher's test was designed like a marking exercise in which teachers had to mark a hypothetical standard four pupil's exam.

Table 29. SDI 2010-2014 Tanzania Trends: Standard four pupils' test scores

	TZ SDI 2010			TZ SDI 2014		
(Percent)	*Tanzania*	*Rural*	*Urban*	*Tanzania*	*Rural*	*Urban*
English *(average score)*	*43*	*41*	*52*	*36*	*31*	*53*
Recognize a letter	70	68	80	72	68	86
Reading a paragraph	27***	24***	45***	36	30	53
Reading comprehension	4***	4***	7***	14	11	24
Mathematics *(average score)*	*39****	*38****	*48****	58	56	65
Adding two double digits	63	60	77	61	58	73
Subtracting two single digits	81	79	87	74	71	85
Subtracting two double digits	37	35	52	39	34	55
Multiplying two double digits	11	10	19	12	10	21
Multiplying two triple digits	7***	6	12***	9	7	17
Dividing two single digits	36	34	51	38	36	44
Dividing double by single digit	14***	12	23	18	16	25
Division task	12	11	14	20	20	17
Multiplication (problem solving)	7	6	11	9	7	14

Note: Levels of significance: *** p<0.01, ** p<0.05, * p<0.1. Source: Tanzania SDI 2010 and 2014 based on author's calculations.

It is puzzling that pupils performed better, while their teachers' performances have been stagnant or might even be declining. This may be partly explained by the fact that pupils have been spending a lot more time with their teachers as shown earlier in this report. This increased teaching time was the resultant of lower school and classroom absence rates, which may have compensated for the slight dip in knowledge. The reduction in class size may also have contributed to improvements in test scores with teachers having more time to spend with individual pupils.

B. Comparing Tanzania to neighbors and others

Table 30 shows the national averages of the indicators for several SDI countries. Bearing in mind that overall performance is low, Tanzania fares relatively well among the SDI countries in terms of inputs, effort, and knowledge. The big-three in the East African Community - Uganda, Kenya, and Tanzania – which aim for stronger regional economic integration, are also included in this comparison. For traditional quality indicators, such as infrastructure, Tanzania is doing worse than its EAC neighbors, but better than Nigeria, Togo, and Mozambique. Ugandan and Kenyan schools, as well as schools in Mozambique, had noticeably better levels of minimum equipment availability (79.5 percent, 74.3 percent, and 76.8 percent, respectively) compared to 62.4 percent for Tanzania. As shown in Table 7, the main constraint was sufficient contrast for pupils to be able to read the blackboard. In contrast, Tanzanian schools fared better than Nigeria (48.2 percent) and Togo (24.3 percent) in regards to minimum equipment availability. A similar picture emerges for infrastructure availability, although the gap in infrastructure between Tanzania and its EAC neighbors was narrower. Tanzania did very poorly with textbook availability, only outperforming Uganda with the average grade four Tanzanian pupil being more than four times more likely to use a textbook in the classroom than her Ugandan peer. Pupils in Togo, Mozambique, Kenya, and Nigeria were, however, more likely to use a textbook in class with 76 percent, 68.1 percent, 44.5 percent, and 33.7 percent of pupils, respectively, doing so.

When it comes to teachers' effort, Tanzania and Kenya displayed patterns with relatively low school absence rates (14.4 percent and 15.2 percent, respectively), but high classroom absence rates with almost half of the teachers not found in classroom at any point in time. Mozambique and Uganda have significantly higher absence rates for both. The three West African countries (Nigeria, Senegal, and

Togo) have much lower classroom absence rates. The difference between school absence rates and classroom absence rates was also narrower, meaning that when teachers went to school they tended to be in the classroom, although they were not necessarily teaching.

Kenyan, Ugandan, and Tanzanian teachers significantly outperformed Nigerian, Togolese, and Mozambican teachers in minimum knowledge. Tanzanian teachers (21.5 percent) are second only to Kenya (34.8 percent), who was the best performer, although it scored modestly with only one out of three teachers scoring above the 80 percent-threshold to be considered as having minimum knowledge of the grade four curriculum. Almost no teachers in Nigeria and Togo passed that bar. Tanzanian teachers' average test scores for all three sections of the skills test (48.3 percent) are also second only to their Kenyan counterparts (55.6 percent). For all countries, teachers performed significantly better in mathematics than language (English for all, except Togo which uses French). There is a similar scenario with the overall pupil test results; once again, Tanzania falls behind Kenya and Uganda with their pupils' test performance (40.1 percent, 45.3 percent, and 69.4 percent, respectively).

Table 30. Comparison of SDI results across countries (public schools only)[24]

	Tanzania 2014	Average SDI	Kenya 2012	Mozambique 2014	Nigeria* 2013	Senegal 2011	Tanzania 2011	Togo 2013	Uganda 2013
Teacher Ability									
Minimum knowledge (At least 80% in language and mathematics)	21.5	12.7	34.8	0.3	2.4	Not Comparable	Not Comparable	0.9	10.1
Test score (language, mathematics, and pedagogy)	48.3	42.0	55.6	26.9	30.5	Not Comparable	Not Comparable	33.9	43.3
Teacher Effort									
School absence rate	14.4	20.1	15.2	44.8	16.9	18.0	23.0	22.6	29.9
Classroom absence rate	46.7	42.1	47.3	56.2	22.8	29.0	53.0	39.3	56.9
Scheduled teaching time	*5h 56min*	*5h 31min*	*5h 31min*	*4h 17min*	*4h 44min*	*4h 36min*	*5h 12min*	*5h 28min*	*7h 13min*
Time spent teaching per day	2h 46min	2h 53min	2h 30min	1h 41 min	3h 10min	3h 15min	2h 04min	3h 15min	2h 56min
Availability of Inputs									
Observed pupil-teacher ratio	43.5	42.1	39.3	21.4	21.5	27.2	52.0	31.4	53.9
Share of pupils with textbooks	25.3	37.2	44.5	68.1	33.7	18.0	19.7	76.0	6.0
Minimum equipment availability (90% with pencils and notebooks)	61.4	57.8	74.3	76.8	48.2	Not Comparable	Not Comparable	24.3	79.5
Minimum infrastructure availability	36.0	36.2	60.2	29.1	13.4	Not Comparable	Not Comparable	14.4	57.2
Pupil Learning									
Test Score (out of 100) (language, mathematics)	40.1	45.4	69.4	20.8	25.1	Not Comparable	Not Comparable	38.1	45.3
Language test score	36.5	44.8	72.5	18.7	23.3	Not Comparable	Not Comparable	36.9	43.4
Mathematics test score	58.2	45.2	57.4	25.1	28.2	Not Comparable	Not Comparable	41.3	41.7

Note: (*) Values for Nigeria are the weighted average of the four states surveyed, namely Anambra, Bauchi, Ekiti, and Niger.

[24] The information in Table 29 is a duplicate of Table 1 in the Executive Summary. It is shown here for ease of reference.

Table 31 provides information on the average scores of the different parts of the test, as well as more details. Tanzania teachers ranked last on language among the five countries shown in Table 31. This is certainly due to the fact that they were tested in English, although they were supposed to have mastered that language. The composition task proved to be the hardest for all teachers, but this was also the area where the gap between Tanzanian teachers and their East African peers was the largest in language test.

In mathematics, Tanzanian teachers were second only to Kenyans. It was mostly for the upper primary part of the test that Tanzanian teachers lost ground. However, they managed to score better than their Kenyan counterparts on comparing fractions. Tanzanian teachers had difficulty in interpreting data on a graph or Venn diagram. For more information on the mathematics results for teachers in Tanzania, Uganda, and Kenya, see Figure C 3. Finally, in pedagogy, Tanzanian teachers fared at par with Kenyans and outperformed Ugandans. They doubled the score of the West African teachers. Tanzanian teachers outscored all others in "preparing a lesson plan" but lost ground on "assessing pupil's abilities."

Table 31. Average teachers' knowledge scores across SDI countries

(Percent)	Tanzania	Kenya	Nigeria	Togo	Uganda
Overall score (language, mathematics, and pedagogy)	**48**	**58**	**38**	**35**	**45**
Language (average score)	***42***	***65***	***49***	***50***	***58***
Grammar task	73	93	64	74	90
Cloze task	53	69	38	30	62
Composition task	21	51	24	26	43
Mathematic (average score)	***63***	***81***	***42***	***33***	***65***
Adding double digit numbers	97	97	89	79	97
Subtracting double digits	86	88	70	65	83
Adding triple digit numbers	85	88	78	60	87
Multiplying two digit numbers	85	87	61	51	76
Adding decimals	64	77	34	36	61
Comparing fractions	81	91	58	15	77
Interpreting a Venn Diagram	49	73	36	22	72
Interpreting data on a graph	27	67	20	14	32
Subtraction of decimal numbers	66	82	45	18	68
One-variable algebra	50	72	15	9	55
Pedagogy (average score)	***36***	***35***	***18***	***19***	***25***
Preparing a lesson plan	58	39	20	27	31
Assessing pupil's abilities	18	33	23	33	25
Evaluating pupils' progress	22	29	6	6	11

Source: Various SDI reports and author's calculations.

VII. CONCLUSION: WHAT DOES THIS MEAN FOR TANZANIA?

Over the past decade, Tanzania has invested a lot of resources in its education sector. According to PMO-RALG (2014), for the fiscal year 2013/14, the education sector accounted for 17.4 percent of the government's expenditure and 6.2 percent of GDP. This is slightly down from 19.1 percent and 6.5 percent, respectively, in 2012/13 but is still in the ballpark of the education budget over the previous decade.

Tanzania earned a return on its education investment with completion rates in primary that increased from 55 percent in 2000 to over 80 percent in 2012. Gender equality has also been achieved in primary education and girls are catching up fast in secondary, although they still lag behind. However, the recent 2012 catastrophic form four exam results were a wake-up call for the Government of Tanzania and all education stakeholders that beyond access indicators, the quality of education is a critical dimension not to be overlooked.

It is now time to invest in quality education and the government has started this process with the Education Big Results Now initiative that clearly recognizes quality as the next frontier and a smart way to improve value-for-money of education public spending. The Service Delivery Indicators (SDI) results show that Tanzania is doing relatively well in terms of teachers' school absence rates, although teachers are not always found in the classroom, even when they are in the school. The SDI also shows that Tanzanian primary schools need some upgrading in terms of infrastructure and teaching equipment, although this is not the most important issue.

With only one out of five teachers mastering the curriculum they teach, teachers' capacity (or its lack thereof) to teach their subject matter is critical and needs to be addressed. There is a wealth of evidence that teachers' knowledge and their capacity to transmit it (i.e., pedagogical skills) is a major determinant of pupils' learning outcomes.

Comparing the 2010 and 2014 SDI surveys, one clearly notices that Tanzania has made substantial progress in some areas, but there are still remaining gaps. Also, the achieved progress is from a very low base and there is still quite a bit of room for the Tanzanian education system to deliver quality to its pupils and get them prepared and equipped to face competition in the national, regional, and international labor markets.

VIII. ANNEXES

ANNEX A: TANZANIA SDI SAMPLING STRATEGY

The overall objective of the SDI is to produce accurate and representative indicators at the national, urban, and rural levels. In some countries, like Tanzania, it may be required that the indicators be representative at a sub-national level (e.g., region or province). The main units of analysis are facilities (schools and health centers) as well as providers (teachers and health workers). In the case of education, the SDI also aims to produce accurate information on standard four pupils' performance on Kiswahili, English, and Mathematics.

A1. Sampling frame for the 2014 Tanzania SDI

The sampling frame for the 2014 Tanzania Education SDI was based on the 2012 EMIS data provided by the Ministry of Education and Vocational Training (MoEVT). The original sample frame contained 15,331 schools with identifier variables, such as region and council. This was close to the 15,362 schools contained in the PLSE 2012 school ranking database. The final sample frame was purged of the six schools that had no standard two pupils, which left the frame with 15,325 primary schools overall. The frame contained only information on standard two's enrollment, whereas SDI was focused on standard four. A separate list of schools, which was used for the distribution of the so-called "radar books," was also obtained. This list had fewer schools (14,120), but enrollment numbers for all grades. That information was used to estimate standard four enrollment. The number of pupils enrolled in standard four was estimated at 1.2 out of a total primary pupil body of 8.1 million.

With 6,025 schools (almost 40 percent) with missing location information (i.e., urban/rural), the sample frame had an important challenge to offer. The issue was dealt with by a two-step procedure. First, all the schools located in a municipal council (MC) or a town council (TC) were considered urban, whereas those in a district council (DC) were tallied in the rural column. This eliminated 1,447 schools, leaving us with 4,578 schools (30 percent) with unknown locations. For the second step, because there was no other variable that could provide information on the location, the remaining schools were randomly split between urban and rural, with 80 percent of the schools considered rural. During the data collection, the head teachers were asked whether their school was urban or rural. This new information was used for post-stratification adjustment.

Although the SDI is usually representative at the national, urban, and rural levels, in Tanzania it was requested that the survey be also representative of the traditional strata in household surveys, which were (1) Dar es Salaam, (2) other urban areas, and (3) rural areas. Because of a large DFID education program, it was agreed to regroup the regions in which the EQUIP-T operates as a single stratum called (4) EQUIP-T. Table A 1 shows the overall sample frame with the number of administrative units, such as councils, the number of standard two pupils (our final variable used for weights), and the total number of primary pupils within each stratum.

Table A 1. 2012 EMIS sample frame by stratum

	# Council	# Schools	# S2 Pupils	# Total Pupils
Dar es Salaam	3	352	69,841	436,952
EQUIP-T	33	3,149	297,576	1,861,751
Other Urban	26	2,097	189,538	1,185,823
Rural	98	9,927	736,937	4,610,562
Total	160	15,325	1,293,892	8,095,088

Source: Author's calculations using MoEVT 2012 EMIS database

The stratification variables provided the domains (strata) and reporting levels (the analysis tables followed these levels) of the survey. The stratification also depended on the most important indicators to be measured in the survey (absence rates and performance levels). Finally, it was advisable to order the clusters within each stratum by variables that were correlated with key survey indicators for further implicit stratification when systematic selection was used.

A multi-stage clustered sampling strategy was adopted for the 2014 Tanzania SDI. The first stage cluster selection was carried out independently within each explicit stratum. The primary cluster considered was the council, which was, therefore, the primary sampling unit (PSU). At the second stage, schools were selected and, at the third stage, teachers and standard four pupils.[25] It was decided that within each stratum, except Dar es Salaam, 25 councils would be chosen with probability proportional to size (number of standard two pupils). Note that this implied, that at this stage, a standard two (and by extension standard four) pupil in each stratum had an equal probability for her council to be selected.

A2. Sample size and sample allocation for the 2014 Tanzania SDI

The optimal sample size of any survey depends on the precision required for the main estimates and resource constraints. The precision of survey estimates depends on the sampling and non-sampling errors. Whereas the sampling error can be measured within a survey, this is not the case for the non-sampling error. The sampling error is smaller the larger the sample, but the non-sampling error grows with the size of the survey. It is, thus, highly advisable to carry out a survey of reasonable sample size that can be managed with effective quality controls to help contain the non-sampling error.

To gauge the precision of the estimate, a previous similar survey or a survey measuring the same indicator is very useful. For Tanzania, a pilot 180-school SDI survey was carried out in 2010. The pilot SDI collected almost identical data to the present survey, therefore, providing us with a very strong advantage for a good measure of design effect and standard errors as basis for the current survey sampling strategy. The design effect is critical for determining the optimal sample size. It is the ratio of the variance of an estimate based on the actual multi-stage sample design and the same variance, if the sample was a simple random one of the same size. The design effect is a measure of the relative efficiency of the sample design.

[25] The selection of teachers and standard four pupils was done once the enumerator was at the school premises. For the purpose of sampling schools, the number of standard two pupils was used as the weight variable with the (reasonable) assumption that the ratio between standard two and standard four pupils was constant.

Table A 2.Teachers' absence rate, average, standard errors, and design effect SDI 2010

	Percent	Std. Err.	[95% Conf. Interval]		Design Effect	Sample Size (Schools)	Sample Size (Teachers)
School absence rate							
Rural	19	1.6	15.9	22.3	2.04	135	1,278
Urban	43	2.9	37.1	48.5	1.94	45	490
Tanzania	27	1.8	23.0	30.3	3.05	180	1,768
Classroom absence rate							
Rural	51	2.3	46.4	55.5	1.60	135	1,278
Urban	59	3.2	52.1	65.0	1.57	45	490
Tanzania	53	1.9	49.6	57.1	2.58	180	1,768

Source: Author's calculations using 2010 SDI data

Table A 2 provides information on teachers' school and classroom absence rates in the 2010 SDI survey, which were estimated at 27 percent and 53 percent, respectively. It also varied a great deal across the urban and rural strata used in the 2010 survey. The design effect for teachers' absence rate was around 3.0 and 2.6, which indicates a more or less efficient sampling strategy (it is, indeed, not uncommon to have design effect above 3.0 for cluster sampling). The standard errors were, however, relatively large, especially for urban areas, as shown by the wide confidence intervals. The 2014 SDI aimed at a national standard error around 1.2 percent for absence rates. Using the 2010 SDI as our basis, it was possible to estimate the necessary sample size, for any given standard error, using the following formula:

$$se_{SDI14}(\bar{a}) \cdot \sqrt{N_{SDI14}} \approx se_{SDI10}(\bar{a}) \cdot \sqrt{N_{SDI10}} \cdot \sqrt{DEFF_{SDI10}/DEFF_{SDI14}}$$

Because the design effect for the 2010 SDI was already at 3.0 for school absence, and the current SDI planned for two more strata, it was expected to keep the design effect around the same level, keeping the last item on the right hand side of the above equation at 1. It was, then, easy to compute the necessary sample size given the objective of a 1.2 standard error. For that standard error, the estimated sample size was 427 schools. Tolerating a slightly higher standard error of 1.3 percent for school absence rate, the sample size came down to 364 schools. It was decided that 400 schools would strike the right balance between the budget and the desired precision.

After determining the sample size, it remained to decide on the sample allocation across strata. Because the number of strata in the 2014 SDI was larger than in the previous survey, we did not use information from 2010 for allocating the 400 schools across the four strata. There are several allocation mechanisms possible for efficient sampling. For the Tanzania 2014 SDI, an adjusted-proportional allocation was used, whereby the share of schools in the stratum was similar to the share of pupils in the stratum compared to the overall population. Adjustments were then made if, for instance, in a given strata the number of schools allocated was too small due to the small pupil population in the stratum. The final sample allocation is given in Table A 3.

Table A 3. SDI sample allocation across regions

	# Schools	# S2 Pupils	Sample allocation
Dar es Salaam	352	69,841	47
EQUIP-T	3,149	297,576	74
Other Urban	2,097	189,538	58
Rural	9,927	736,937	221
Total	15,325	1,293,892	400

A3. Sampling schools, teachers, and pupils

Now that the total sample size and its allocation across strata had been decided, it remained to sample the actual schools that were included in the final sample and, within each school, the pupils and teachers to be assessed. This was done using a two-stage sampling method. First, in each stratum schools were chosen within the selected councils. Once at a selected school, the enumerator selected teachers and pupils depending on the structure of the classrooms.

The schools were chosen using probability proportional to size (PPS), where size was the number of standard two pupils as provided by the 2012 EMIS database. As for the selection of the cluster, the use of PPS implied that each standard four pupil within a stratum had an equal probability for her school to be selected.

Finally, within each school, up to 10 standard four pupils and 10 teachers were selected. Pupils were randomly selected among the standard four pupil body, whereas for teachers, there were two different procedures for measuring absence rate and assessing knowledge. For absence rate, 10 teachers were randomly selected from the teachers' roster and the whereabouts of those teachers was ascertained in a return surprise visit. For the knowledge assessment, however, all teachers who were currently teaching in primary four or taught primary three the previous school year were included in the sample. Then a random number of teachers in upper grades were included to top up the sample. These procedures implied that pupils across strata, as well as teachers across strata and within a school (for the knowledge assessment) did not all have the same probability of selection. It was, therefore, warranted to compute weights for reporting the survey results.

A4. Weights for schools, teachers, and pupils

To be representative of the population of interest, sample estimates from the 2014 Tanzania SDI had to be properly weighted, using a sampling weight, or expansion factor. Note that different weights needed to be applied depending on the relevant level for the variable, which could be the school, teacher, or pupil. The basic weight for each entity was equal to the inverse of its probability of selection, which was computed by multiplying the probabilities of selection at each sampling stage. All the weights were computed and included in the dataset.

ANNEX B: DEFINITION OF INDICATORS

School absence rate	
Share of a maximum of 10 randomly selected teachers absent from school during an unannounced visit	It is measured as the share of teachers who are absent from school at a time of an unannounced visit. It is measured in the following way: During the first announced visit, a maximum of ten teachers are randomly selected from the list of all teachers (excludes volunteer and part time teachers) who are on the school roster. The whereabouts of these ten teachers are then verified in the second, unannounced, visit. Teachers found anywhere on the school premises are marked as present.
Classroom absence rate	
Share of teachers who are present in the classroom during scheduled teaching hours as observed during an unannounced visit	The indicator is measured as the share of teachers not in the classroom at the time of an unannounced visit. The indicator is constructed in the same way as school absence rate indicator, with the exception that the numerator now is the number of teachers who are either absent from school, or present at school but absent from the classroom.
Time spent teaching per day	
Amount of time a teacher spends teaching during a school day	This indicator reflects the typical time that teachers spends teaching on an average day. This indicator combines data from the staff roster module (used to measure absence rate), the classroom observation module, and reported teaching hours. The teaching time is adjusted for the time teachers are absent from the classroom, on average, and for the time the teacher teaches while in classrooms based on classroom observations. While inside the classroom distinction is made between teaching and non-teaching activities. Teaching is defined very broadly, including actively interacting with students, correcting or grading student's work, asking questions, testing, using the blackboard or having students working on a specific task, drilling or memorization. Non-teaching activities includes working on private matters, maintaining discipline in class or doing nothing and thus leaving students not paying attention.
Minimum knowledge	
Share of teachers with minimum knowledge	It is measured as the percentage of teachers who can master the curriculum they taught. It is based on a mathematics and language tests covering the primary curriculum administered at the school and is calculated as the percentage of teacher who score more than 80% on the language and mathematics portion of the test. The test is given to all mathematics or language teachers that taught 3rd grade last year or 4th grade in the year the survey was conducted.
Test score	It is measured as the overall score of a mathematics, language and pedagogy tests covering the primary curriculum administered at the school level to all mathematics and language teachers that taught 3rd grade last year or 4th grade in the year the survey was conducted.

Minimum infrastructure availability	
Unweighted average of the proportion of schools with the following available: functioning electricity and sanitation	It is a binary indicator capturing availability of: (i) functioning toilets and (ii) classroom visibility. Functioning toilets is defined as whether toilets were functioning, accessible, clean and private (enclosed and with gender separation) as verified by an enumerator. To verify classroom visibility we randomly select one 4th grade classroom in which the enumerator places a printout on the board and checks whether it was possible to read the printout from the back of the classroom
Minimum equipment availability	
Unweighted average of the proportion of schools with the following available: functioning blackboard with chalk, pens or pencils, and notebooks or paper	It is a binary indicator capturing availability of: (i) functioning blackboard and chalk and (ii) pens, pencils and exercise books9 in 4th grade classrooms. In one randomly selected 4th grade classroom in the school the enumerator assessed if there was a functioning blackboard by looking at whether text written on the blackboard could be read at the front and back of the classroom, and whether there was chalk available to write on the blackboard. We considered that the classroom meet the minimum requirement of pens, pencils and exercise books if both the share of students with pen or pencils and the share of students with exercise books are above 90%.
Share of pupils with textbooks	
Number of mathematics and language books used in a grade four classroom divided by the number of pupils present in the classroom	The indicator reflect the typical ratio in student to textbooks in the 4th grade classroom. It is measured as the number of students with the relevant textbooks (mathematic or language conditional on which randomly selected class is observed) in one randomly selected 4th grade class and divided by the number of students in that classroom.
Observed pupil-teacher ratio	
Average number of grade four pupils per grade four teacher	This indicator reflects the typical ratio in pupils to teachers in the 4th grade classroom. It is measured as the number of students in one randomly selected 4th grade class at the school.

ANNEX C: ADDITIONAL RESULTS

Table C 1. School Inputs

	Tanzania	Dar es Salaam	Other Urban	Rural	EQUIP-T	Rural	Urban
Pupils per teacher (units)	**43.7**	**38.4**	**37.4**	**44.7**	**52.6**	**44.7**	**37.3**
Observed pupil-teacher ratio	**43.5**	**69.8**	**58.9**	**40.7**	**46.1**	**40.8**	**60.4**
Share of pupils with pencils	95.8	96.9	96.7	95.7	97.0	95.7	96.7
Share of pupils with paper	96.3	96.8	97.3	96.2	97.5	96.2	97.2
Have a board (% of classrooms)	98.4	99.1	97.2	98.5	94.4	98.5	97.5
Have chalk (% of classrooms)	97.0	100.0	96.2	97.1	100.0	97.1	96.8
Sufficient contrast to read board (% of classrooms)	73.9	93.1	88.4	71.4	66.2	71.4	89.0
Minimum equipment availability (% of classrooms)	**61.4**	**83.2**	**80.1**	**58.3**	**58.9**	**58.3**	**80.4**
Share of pupils with textbooks	**25.3**	**31.0**	**14.1**	**26.7**	**22.2**	**26.7**	**16.7**
Share pupils with mathematics textbooks	24.6	28.4	12.2	26.2	19.8	26.3	14.7
Share pupils with English textbooks	26.3	34.9	16.8	27.4	25.9	27.4	19.7
Functioning toilet (% of schools)	47.0	73.0	69.5	43.2	42.2	43.3	70.1
Has toilet (% of schools)	97.1	98.8	100.0	96.6	92.9	96.6	99.8
Toilet clean (% of schools)	92.0	95.3	96.1	91.4	90.7	91.4	95.9
Toilet private (% of schools)	56.2	83.3	79.7	52.2	53.4	52.3	80.6
Toilet accessible (% of schools)	96.6	94.4	92.9	97.2	98.0	97.2	93.1
Visibility judged by enumerator (% of classrooms)	75.8	94.0	89.4	73.4	71.8	73.5	90.1
Minimum infrastructure availability (% of schools)	**40.4**	**67.0**	**61.4**	**36.8**	**34.3**	**36.9**	**62.3**

Figure C 1. Regional distribution of teachers' school and classroom absence rates

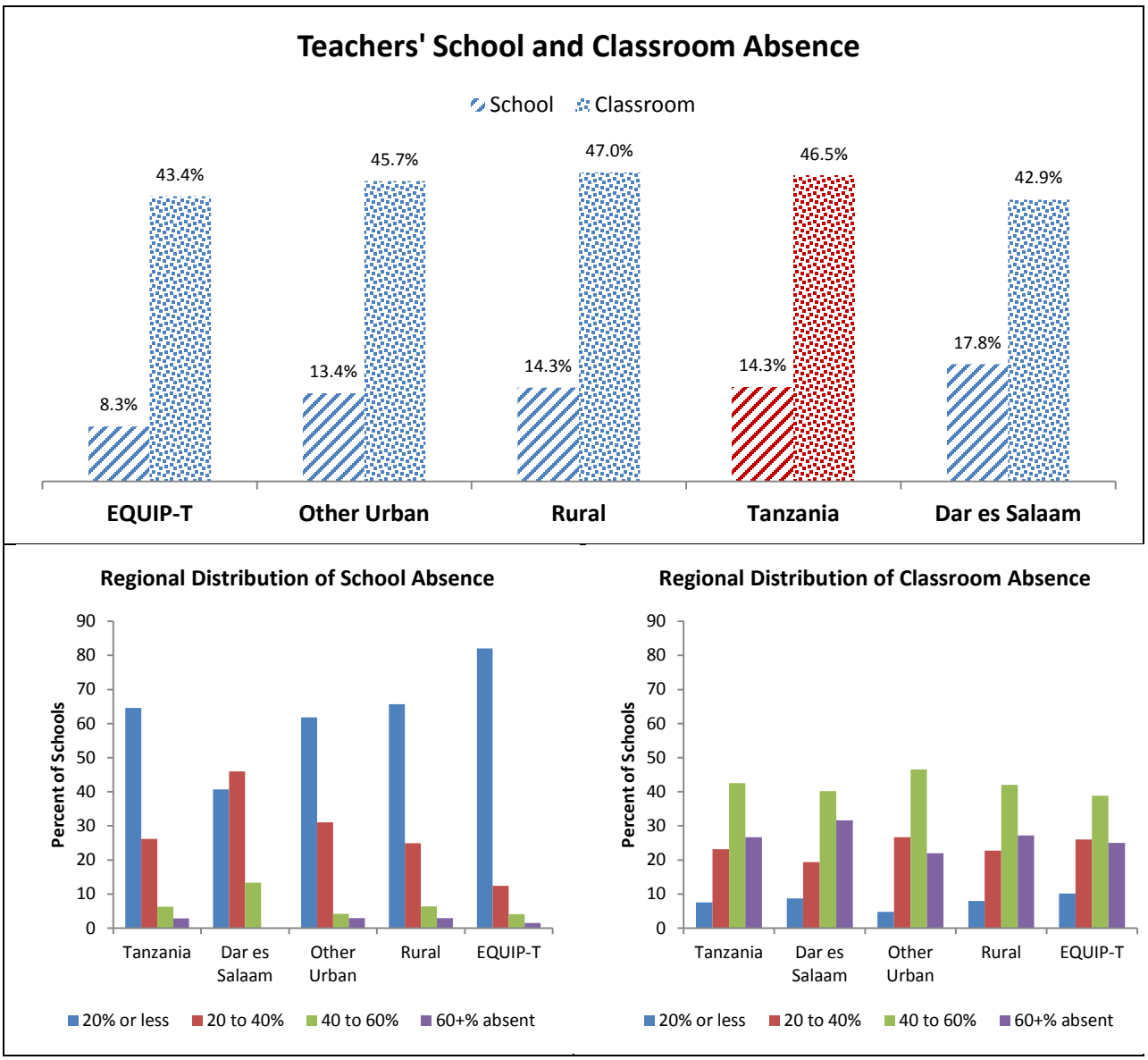


Table C 2. Official teaching time, loss of teaching time, and time spent teaching per day

	Scheduled teaching time (in minutes)				Share of time devoted to teaching activities				Time spent teaching per day (in minutes)			
	Average time	Robust Std. Err.	[95% Conf. Inter.]		Percent	Robust Std. Err.	[95% Conf. Inter.]		Average time	Robust Std. Err.	[95% Conf. Inter.]	
Tanzania	355.4	3.0	349.4	361.3	89.0	1.3	86.4	91.5	167.5	13.0	142.0	193.0
Dar es Salaam	341.0	8.1	325.1	356.9	86.6	4.4	78.0	95.2	186.6	30.3	127.1	246.2
Other Urban	348.3	5.8	336.8	359.8	86.6	2.6	81.4	91.8	156.3	24.1	108.9	203.7
Rural	356.8	3.4	350.1	363.4	89.4	1.5	86.5	92.2	168.4	14.7	139.5	197.4
EQUIP-T	358.1	5.9	346.3	369.9	80.9	4.5	71.9	89.8	187.0	25.5	136.2	237.7
Rural	356.7	3.4	350.0	363.4	89.4	1.4	86.5	92.2	168.7	14.7	139.8	197.6
Urban	347.3	5.1	337.4	357.3	86.4	2.3	81.8	91.0	160.0	20.9	118.9	201.0

Source: Tanzania SDI 2014 and author's calculations

Figure C 2. Orphan classrooms

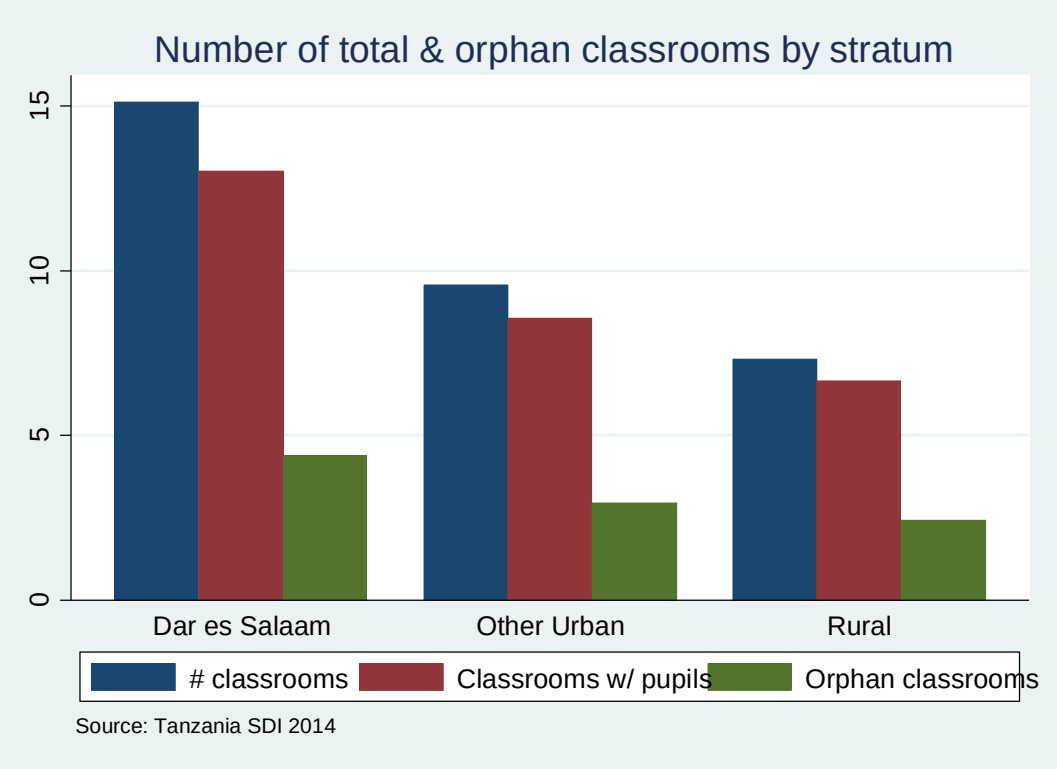


Table C 3. Teachers' mathematics assessment scores

(Percent)	Tanzania	Dar es Salaam	Other Urban	Rural	EQUIP-T	Rural	Urban
Mathematics (complete test)	63.1	69.1	64.6	62.5	59.3	62.5	65.4
Lower Primary	69.8	75.9	71.7	69.2	66.9	69.2	72.5
Upper Primary	50.6	56.3	51.5	50.1	44.9	50.1	52.3
Adding double digit numbers	97.0	99.1	97.3	96.8	95.1	96.8	97.6
Subtracting double digit numbers	85.8	87.9	85.2	85.8	81.9	85.8	85.5
Adding triple digit numbers	85.3	91.2	89.5	84.1	86.2	84.1	90.0
Dividing double by single	80.6	78.1	77.0	81.5	75.6	81.5	77.3
Multiplying two digit numbers	84.9	85.8	85.1	84.9	84.0	84.9	85.3
Adding decimals	64.0	63.9	66.5	63.4	62.3	63.5	66.0
Division two-digit numbers - conceptual understanding	81.3	85.1	86.8	80.0	81.2	80.0	86.6
Comparing fractions different denominators	49.6	65.2	53.7	48.0	46.9	48.0	55.7
Monetary units - multiplication	55.0	61.7	55.0	54.7	55.2	54.7	56.3
Geometry - 2D shapes	91.1	99.3	89.4	91.1	89.6	91.2	91.0
Geometry - types of lines	75.8	81.7	81.6	74.3	70.3	74.3	81.4
Time (reading a clock) - problem solving	47.5	55.8	55.3	45.4	40.7	45.3	55.7
Interpreting data on a Venn diagram	48.9	58.4	49.8	48.3	42.6	48.3	51.2
Interpreting data on a graph	27.5	24.4	25.5	28.1	23.8	28.0	25.5
Square root (no remainder)	79.3	92.6	78.5	78.9	73.7	78.9	80.8
Subtraction of numbers with decimals	66.5	74.5	66.6	66.1	62.7	66.2	67.8
Division of fractions	58.3	63.0	58.6	58.0	53.6	58.0	59.4
One variable algebra	50.5	53.6	50.9	50.2	44.2	50.2	51.3
Geometry - computing perimeter of a rectangle	51.3	63.3	55.4	49.8	42.4	49.8	56.8
Geometry - computing area of a rectangle	46.1	56.5	50.2	44.7	38.4	44.7	51.5

Source: Tanzania SDI 2014 and author's calculations

Table C 4. Pupils' test scores – Language

(Percent)	Tanzania	Dar es Salaam	Other Urban	Rural	EQUIP-T	Rural	Urban	Boy	Girl	No breakfast	Breakfast
Average Scores											
English & Mathematics	40.1	65.0	52.1	35.2	32.6	35.5	55.0	41.7	39.1	40.5	40.0
Kisawhili & Mathematics	76.2	92.6	84.8	72.8	69.3	73.0	86.5	77.8	74.7	75.1	77.9
English	36.5	63.2	50.3	31.1	29.0	31.4	53.3	38.0	35.6	36.9	36.5
Mathematics	80.9	97.5	90.4	77.3	74.0	77.5	91.9	82.5	79.4	79.5	83.1
Kiswahili	58.2	75.0	62.1	55.9	50.6	56.1	65.0	60.1	56.9	58.5	58.2
Non-verbal reasoning	53.6	67.7	57.4	51.5	47.3	51.8	59.2	55.3	52.4	54.6	52.3
English test-takers											
Can read a letter	72.3	93.5	83.3	67.9	70.1	68.2	85.6	72.9	71.7	74.0	69.4
Can read a word	73.5	92.0	84.7	69.4	70.0	69.6	86.4	73.8	73.3	74.8	71.3
Has basic vocabulary	30.8	64.1	48.4	23.9	20.6	24.2	52.2	33.5	28.2	31.4	29.7
Can read a sentence	30.9	63.4	47.2	24.4	20.8	24.7	51.1	32.8	29.1	31.0	30.6
Can read a paragraph	2.3	8.9	5.0	1.1	0.6	1.1	6.1	2.7	1.9	2.1	2.5
Comprehension (factual)	12.4	39.6	16.4	9.2	4.6	9.6	21.6	11.7	13.0	13.9	9.8
Comprehension (analytic)	17.1	33.9	20.8	14.9	7.6	15.0	24.0	18.3	16.0	16.6	17.9
Kiswahili test-takers											
Can read a letter	86.7	95.8	91.2	84.9	76.3	85.1	92.1	88.0	85.7	84.3	90.6
Can read a word	88.7	98.8	95.4	86.3	83.9	86.4	96.1	89.6	87.9	87.5	90.7
Has basic vocabulary	85.1	97.3	90.2	82.9	75.5	83.0	91.7	85.3	84.9	83.9	87.0
Can read a sentence	83.0	98.7	92.5	79.5	77.1	79.7	93.8	84.4	81.8	81.0	86.3
Can read a paragraph	74.8	97.4	84.1	70.7	62.6	71.0	87.0	78.2	71.8	72.9	77.8
Comprehension (factual)	63.2	93.1	74.2	58.2	42.0	58.5	78.7	63.2	63.3	61.6	65.9
Comprehension (analytic)	42.7	70.3	50.4	38.6	30.8	39.1	54.3	40.3	44.7	43.5	41.3

Table C 5. Pupils' test scores – Mathematics

(Percent)	Tanzania	Dar es Salaam	Other Urban	Rural	EQUIP-T	Rural	Urban	Boy	Girl	No breakfast	Breakfast
Average Scores											
English & mathematics	40.1	65.0	52.1	35.2	32.6	35.5	55.0	41.7	39.1	40.5	40.0
Kisawhili & mathematics	76.2	92.6	84.8	72.8	69.3	73.0	86.5	77.8	74.7	75.1	77.9
Mathematics	80.9	97.5	90.4	77.3	74.0	77.5	91.9	82.5	79.4	79.5	83.1
Non-verbal reasoning	53.6	67.7	57.4	51.5	47.3	51.8	59.2	55.3	52.4	54.6	52.3
All test-takers											
Number recognition	95.6	99.0	98.1	94.7	90.7	94.8	98.3	95.5	95.7	94.9	96.9
Ordering numbers	45.1	68.4	48.3	42.4	33.3	42.8	52.4	48.0	42.3	46.1	43.3
Addition (one-digit)	79.7	94.5	85.8	77.0	72.9	77.2	87.6	80.4	78.9	79.0	80.7
Addition (two-digit)	61.1	89.2	68.0	57.2	53.6	57.6	72.5	63.5	59.0	61.0	61.3
Addition (three-digit)	60.3	89.2	69.5	55.7	49.9	56.0	73.9	63.2	57.5	60.0	60.6
Subtraction (one-digit)	74.1	91.6	82.5	70.7	66.8	70.8	84.9	76.4	72.0	73.7	74.9
Subtraction (two-digit)	39.3	69.7	50.7	34.1	31.1	34.5	55.0	41.3	37.5	40.5	37.2
Multiplication (one-digit)	37.6	55.4	37.7	36.2	27.0	36.5	41.4	40.4	35.0	38.1	36.7
Multiplication (two-digit)	12.1	34.0	16.8	9.2	7.7	9.5	20.6	14.8	9.6	12.0	12.3
Multiplication (three-digit)	9.2	27.9	13.1	6.8	7.7	7.0	16.6	10.6	8.0	9.6	8.6
Division (one-digit)	37.9	56.6	39.2	36.1	29.2	36.1	43.5	40.0	36.0	39.7	34.7
Division (two-digit)	18.1	40.5	19.7	15.9	13.3	16.1	24.6	20.4	15.9	19.2	16.1
Division (analytical)	19.6	24.9	14.6	20.4	16.1	20.4	17.1	21.3	18.0	18.8	20.9
Multiplication (prb. solv.)	8.9	23.8	10.1	7.4	4.2	7.4	13.6	10.4	7.4	9.9	7.2
Complete sequence	14.1	26.9	13.2	13.3	6.8	13.3	16.6	15.2	13.0	15.4	11.8

Table C 6. Correlations between the SDI and test scores – English

	School absence rate	Classroom absence rate	Time spent teaching per day	Share of teachers with minimum knowledge	Teacher test score (English)	Teacher test score (Maths)	Minimum equipment availability	Minimum infrastructure availability	Observed pupil-teacher ratio	Share of pupils with textbooks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A – All Schools										
Coef.	0.0132	-0.0323	-0.0497*	0.126**	0.0695	0.106**	0.141**	0.336***	-0.169	-0.0237
Std. Err	(0.0654)	(0.0388)	(0.0175)	(0.0381)	(0.112)	(0.0252)	(0.0382)	(0.0486)	(0.0941)	(0.0230)
Observations	396	396	394	395	395	395	396	396	396	395
Adj. R-square	0.000	0.001	0.003	0.018	0.006	0.013	0.023	0.116	0.029	0.001
Panel B – Urban Schools										
Coef.	-0.0886	-0.118	0.0818	0.147	0.300**	0.0628	-0.0498	0.252**	-0.346**	0.221
Std. Err	(0.206)	(0.111)	(0.109)	(0.158)	(0.0460)	(0.100)	(0.0802)	(0.0405)	(0.0613)	(0.105)
Observations	110	110	110	110	110	110	110	110	110	110
Adj. R-square	0.008	0.013	0.008	0.020	0.093	0.004	0.002	0.081	0.080	0.043
Panel C- Rural Schools										
Coef.	0.0141	-0.0135	-0.0611***	0.110	0.00339	0.102**	0.102***	0.283**	-0.0958	-0.00545
Std. Err	(0.0558)	(0.0338)	(0.00597)	(0.0438)	(0.0611)	(0.0164)	(0.00767)	(0.0578)	(0.0430)	(0.00214)
Observations	286	286	284	285	285	285	286	286	286	285
Adj. R-square	0.000	0.000	0.005	0.017	0.000	0.015	0.014	0.092	0.011	0.000

Note: Each cell represent a regression where test score is regressed on the indicator noted in the column and a constant. The regression uses sampling weights. Panel A is all schools. Panel B is public schools, controlling for rural-urban location. Weighted robust standard errors in parenthesis. Time spent teaching is measured in minutes. Levels of significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C 7. Correlations between the SDI and test scores – Kiswahili

	School absence rate	Classroom absence rate	Time spent teaching per day	Share of teachers with minimum knowledge	Teacher test score (English)	Teacher test score (Maths)	Minimum equipment availability	Minimum infrastructure availability	Observed pupil-teacher ratio	Share of pupils with textbooks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A – All Schools										
Coef.	-0.0774*	-0.0830**	-0.249**	0.168	-0.0842	-0.0617	0.142**	0.348***	-0.257***	0.104***
Std. Err	(0.0250)	(0.0185)	(0.0739)	(0.0928)	(0.0788)	(0.0313)	(0.0263)	(0.0234)	(0.0422)	(0.0142)
Observations	396	396	394	395	395	395	396	396	396	395
Adj. R-square	0.005	0.005	0.046	0.022	0.005	0.003	0.016	0.084	0.045	0.009
Panel B – Urban Schools										
Coef.	-0.154	-0.0348	2.76e-05	0.337	0.0328	0.151*	-0.188*	0.206*	-0.360	0.199**
Std. Err	(0.137)	(0.0579)	(0.117)	(0.155)	(0.142)	(0.0466)	(0.0463)	(0.0584)	(0.192)	(0.0240)
Observations	110	110	110	110	110	110	110	110	110	110
Adj. R-square	0.026	0.001	0.000	0.115	0.001	0.024	0.037	0.059	0.093	0.038
Panel C- Rural Schools										
Coef.	-0.0745**	-0.0836	-0.281**	0.144	-0.119	-0.0874**	0.149**	0.344***	-0.223***	0.118***
Std. Err	(0.0169)	(0.0287)	(0.0625)	(0.116)	(0.0570)	(0.0131)	(0.0210)	(0.0314)	(0.00887)	(0.000600)
Observations	286	286	284	285	285	285	286	286	286	285
Adj. R-square	0.005	0.005	0.056	0.016	0.011	0.006	0.017	0.075	0.034	0.012

Notes: Each cell represent a regression where test score is regressed on the indicator noted in the column and a constant. The regression uses sampling weights. Panel A is all schools. Panel B is public schools, controlling for rural-urban location. Weighted robust standard errors in parenthesis. Time spent teaching is measured in minutes. Levels of significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C 8. Correlations between the SDI and test scores – Mathematics

	School absence rate	Classroom absence rate	Time spent teaching per day	Share of teachers with minimum	Teacher test score (English)	Teacher test score (Maths)	Minimum equipment availability	Minimum infrastructure availability	Observed pupil-teacher ratio	Share of pupils with textbooks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A – All Schools										
Coef.	0.0399	-0.0851	-0.0115	0.140	0.103	0.194***	0.0290	0.210***	-0.129*	0.0591***
Std. Err	(0.0476)	(0.0454)	(0.0444)	(0.0800)	(0.112)	(0.0139)	(0.0334)	(0.0325)	(0.0483)	(0.00929)
Observations	396	396	394	395	395	395	396	396	396	395
Adj. R-square	0.002	0.007	0.000	0.021	0.011	0.040	0.001	0.042	0.015	0.004
Panel B – Urban Schools										
Coef.	0.00978	-0.0566	0.0196	0.236	0.218**	0.144**	-0.175**	0.0926*	-0.257	0.289*
Std. Err	(0.0751)	(0.0661)	(0.0660)	(0.171)	(0.0460)	(0.0324)	(0.0393)	(0.0307)	(0.0935)	(0.0950)
Observations	110	110	110	110	110	110	110	110	110	110
Adj. R-square	0.000	0.003	0.001	0.058	0.055	0.022	0.033	0.012	0.049	0.083
Panel C- Rural Schools										
Coef.	0.0372	-0.0826	-0.0115	0.122	0.0663	0.195***	0.0131	0.189***	-0.0841	0.0645***
Std. Err	(0.0420)	(0.0466)	(0.0527)	(0.107)	(0.0984)	(0.00867)	(0.0229)	(0.0149)	(0.0574)	(0.00132)
Observations	286	286	284	285	285	285	286	286	286	285
Adj. R-square	0.002	0.007	0.000	0.017	0.005	0.043	0.000	0.033	0.007	0.005

Notes: Each cell represent a regression where test score is regressed on the indicator noted in the column and a constant. The regression uses sampling weights. Panel A is all schools. Panel B is public schools, controlling for rural-urban location. Weighted robust standard errors in parenthesis. Time spent teaching is measured in minutes. Levels of significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C 9. Correlations between the SDI and test scores – Overall score with Kiswahili as language

	School absence rate	Classroom absence rate	Time spent teaching per day	Share of teachers with minimum	Teacher test score (English)	Teacher test score (Maths)	Minimum equipment availability	Minimum infrastructure availability	Observed pupil-teacher ratio	Share of pupils with textbooks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A – All Schools										
Coef.	-0.0320	-0.0735***	-0.168*	0.170	-0.0420	0.0289	0.114***	0.294***	-0.228**	0.0649**
Std. Err	(0.0333)	(0.0116)	(0.0670)	(0.0883)	(0.0772)	(0.0262)	(0.0193)	(0.0176)	(0.0466)	(0.0136)
Observations	396	396	394	395	395	395	396	396	396	395
Adj. R-square	0.001	0.004	0.024	0.026	0.002	0.001	0.012	0.070	0.041	0.004
Panel B – Urban Schools										
Coef.	-0.0737	-0.0681	0.0217	0.276	0.0391	0.173**	-0.160	0.147	-0.308	0.221**
Std. Err	(0.167)	(0.0690)	(0.0763)	(0.113)	(0.131)	(0.0372)	(0.0839)	(0.0654)	(0.195)	(0.0476)
Observations	110	110	110	110	110	110	110	110	110	110
Adj. R-square	0.007	0.005	0.001	0.087	0.002	0.035	0.030	0.033	0.077	0.053
Panel C- Rural Schools										
Coef.	-0.0337	-0.0684**	-0.190*	0.151	-0.0768	0.00928	0.107**	0.280***	-0.188**	0.0767**
Std. Err	(0.0201)	(0.0112)	(0.0646)	(0.109)	(0.0552)	(0.0148)	(0.0167)	(0.0141)	(0.0269)	(0.0162)
Observations	286	286	284	285	285	285	286	286	286	285
Adj. R-square	0.001	0.004	0.030	0.021	0.005	0.000	0.010	0.059	0.028	0.006

Notes: Each cell represent a regression where test score is regressed on the indicator noted in the column and a constant. The regression uses sampling weights. Panel A is all schools. Panel B is public schools, controlling for rural-urban location. Weighted robust standard errors in parenthesis. Time spent teaching is measured in minutes. Levels of significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C 10. Correlations between the SDI and test scores – Overall score with English as language

	School absence rate	Classroom absence rate	Time spent teaching per day	Share of teachers with minimum knowledge	Teacher test score (English)	Teacher test score (Maths)	Minimum equipment availability	Minimum infrastructure availability	Observed pupil-teacher ratio	Share of pupils with textbooks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A – All Schools										
Coef.	0.0257	-0.0636**	-0.0350	0.140*	0.102	0.157***	0.100*	0.313***	-0.160	0.0196
Std. Err	(0.0633)	(0.0192)	(0.0253)	(0.0571)	(0.128)	(0.0219)	(0.0378)	(0.0371)	(0.0690)	(0.0212)
Observations	396	396	394	395	395	395	396	396	396	395
Adj. R-square	0.001	0.004	0.001	0.022	0.012	0.028	0.011	0.099	0.025	0.000
Panel B – Urban Schools										
Coef.	-0.0642	-0.0925	0.0557	0.206	0.321***	0.0923	-0.118	0.212**	-0.345*	0.275
Std. Err	(0.147)	(0.0597)	(0.106)	(0.175)	(0.0302)	(0.0693)	(0.0584)	(0.0262)	(0.104)	(0.108)
Observations	110	110	110	110	110	110	110	110	110	110
Adj. R-square	0.004	0.008	0.004	0.040	0.106	0.008	0.014	0.057	0.079	0.067
Panel C- Rural Schools										
Coef.	0.0263	-0.0521***	-0.0414	0.121	0.0419	0.156***	0.0698**	0.270**	-0.0934***	0.0333**
Std. Err	(0.0576)	(0.00264)	(0.0236)	(0.0757)	(0.0895)	(0.0138)	(0.0122)	(0.0311)	(0.00890)	(0.00697)
Observations	286	286	284	285	285	285	286	286	286	285
Adj. R-square	0.001	0.003	0.002	0.019	0.002	0.032	0.006	0.079	0.010	0.002

Notes: Each cell represent a regression where test score is regressed on the indicator noted in the column and a constant. The regression uses sampling weights. Panel A is all schools. Panel B is public schools, controlling for rural-urban location. Weighted robust standard errors in parenthesis. Time spent teaching is measured in minutes. Levels of significance: *** p<0.01, ** p<0.05, * p<0.1.

Table C 11. Teaching practices by gender

	Male Teacher				Female Teacher			
	Percent	Std. Err.	[95% Conf. Interval]		Percent	Std. Err.	[95% Conf. Interval]	
Teacher used textbook	85.3	2.5	80.3	90.3	89.5	2.3	85.0	93.9
Teacher wrote on blackboard	99.8	0.3	99.2	100.4	99.3	0.6	98.0	100.5
Pupils wrote on blackboard	43.4	3.5	36.4	50.3	50.8	3.7	43.5	58.1
Teacher visited pupils	54.0	3.6	47.0	61.0	59.6	3.6	52.5	66.8
Called pupils by name	78.4	2.9	72.6	84.1	86.6	2.5	81.6	91.6
Teacher was smiling/joking	65.1	3.4	58.5	71.8	57.5	3.7	50.3	64.7
Teacher was hitting/scolding	3.1	1.2	0.7	5.6	5.0	1.6	1.8	8.1
Asked to apply new info.	69.5	3.3	63.0	75.9	67.5	3.5	60.7	74.3
Tested creativity	54.8	3.5	47.8	61.8	68.7	3.4	62.0	75.5
Gave positive feedback	70.0	3.3	63.5	76.4	82.8	2.8	77.3	88.3
Gave corrective feedback	80.6	2.8	75.0	86.2	83.2	2.8	77.7	88.7
Introduced lesson	86.3	2.4	81.5	91.1	89.7	2.3	85.2	94.1
Summarized lesson	42.7	3.5	35.7	49.6	52.6	3.7	45.3	59.9
Assigned homework	40.0	3.5	33.1	46.8	44.1	3.7	36.9	51.4
Reviewed homework	21.6	2.9	15.8	27.4	19.1	2.9	13.3	24.8
Used local language	2.3	1.1	0.2	4.4	1.0	0.8	-0.4	2.5

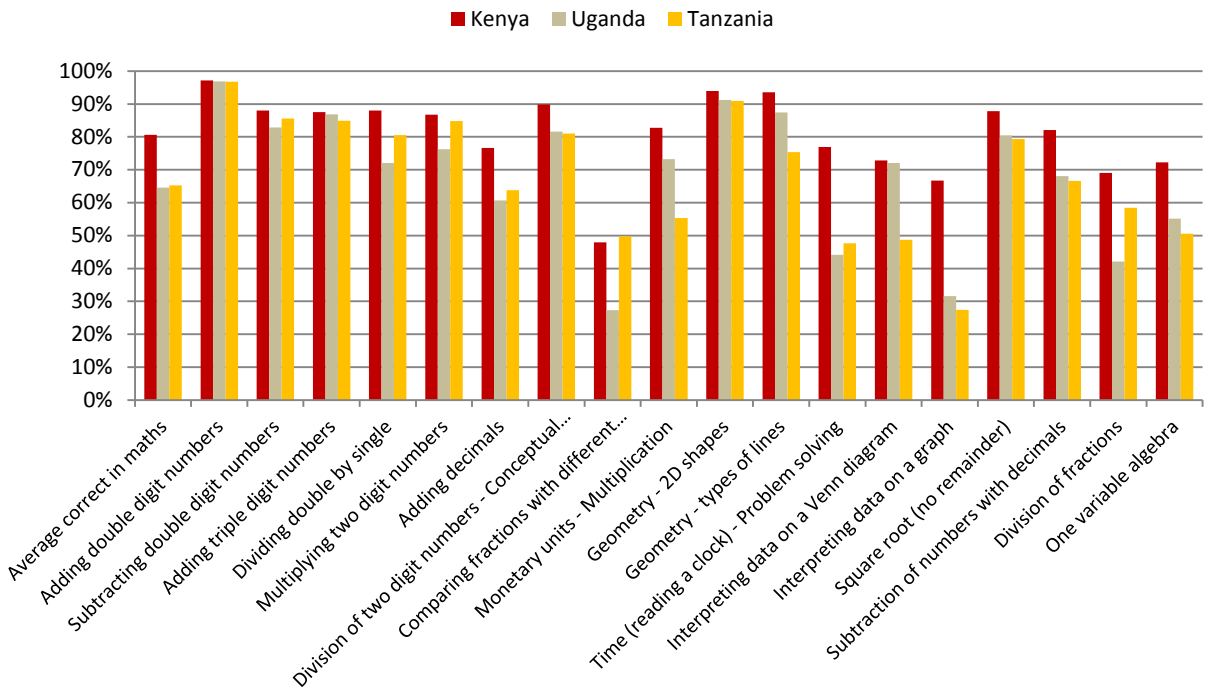
Source: Author's calculations using 2014 Tanzania SDI data.

Table C 12. Regression results by specific classroom teaching practices

	HF-TM	HM-TF	HF-TF	% female teachers	Other urban	Rural	# Obs.	R-squared
# pupils in classroom	-2.097	2.513*	-4.099*	16.58	-4.236**	-22.44**	396	0.151
% of girls in classroom	-0.0251	0.00840	0.00866	0.0174	0.00937**	0.0303**	396	0.013
Share pupils have book	0.138*	-0.0370**	0.0221	0.159***	-0.129***	0.0183	395	0.042
% of girls with textbook	0.0934	0.0573**	0.0870**	0.121**	0.0320***	0.0852***	233	0.063
Share pupils used textbook	0.0762***	-0.0197**	0.131*	0.0156	-0.149***	0.000914	392	0.029
% girls using textbook	0.0234	0.00927	0.0169	0.108	0.0348*	0.0212	214	0.051
Teacher wrote on board	0.0111	0.00392	-0.00286	-0.0104	-0.0161***	-0.015***	394	0.003
Pupils wrote on board	-0.115	0.106**	0.145	-0.317	-0.00345	-0.00916	394	0.022
% girls wrote on board	-0.0637	-0.100	-0.088***	0.347**	0.135***	0.0872*	170	0.065
Share of pupils with pens/pencils	0.0263	-0.0179	-0.0118	-0.00984	0.00152	-0.0124*	393	0.010
% girls with pen/pencil	-0.0211	0.00637	0.0178**	0.0238	0.0254***	0.0564**	392	0.024
Share of pupils with exercise books	0.0408	0.00218	-0.00267	-0.00778	0.0126***	0.00864	393	0.010
% girls with exercise book	-0.0286	0.00493	0.00735	0.0218	0.0174***	0.0484**	393	0.021
Teacher went to pupils	0.0433	-0.0201	0.0692	0.0704	0.0400**	-0.0126	393	0.010
% girls teacher went to	-0.090***	0.0550**	0.0505	-0.00233	-0.0332***	0.0109	209	0.075
Teacher called pupils' name	0.152	0.0387	0.0149	-0.0382	0.0538***	-0.00914	393	0.011
% of pupils teacher called	0.0175	-0.0139**	0.0124	-0.0884*	-0.0131**	0.00466	319	0.045
Teacher kept attendance	0.0259	0.0410	0.0468	0.0246	0.0149	0.0693	396	0.008
Teacher had scheme of work	0.0106	-0.0142	-0.0271	0.199**	0.107***	0.0922**	396	0.012
Teacher had lesson plan	-0.153	-0.0707	-0.0681	0.394**	0.00293	-0.00211	396	0.050
Teacher introduced lesson	-0.0515	0.00399	-0.0167	-0.0657	-0.0177**	-0.113***	389	0.013
Teacher summarized lesson	0.0116	0.0889	0.0267	-0.102	0.146***	0.0169	393	0.014
Teacher assigned homework	-0.150	0.0848**	0.186	-0.315*	-0.173***	-0.0677	393	0.038
Teacher reviewed homework	0.0213	0.0616	0.0870	-0.157*	-0.0149**	-0.00179	382	0.007
Teacher hit pupils	0.0728	0.0237	0.0149	-0.0195	-0.0183**	-0.0187	394	0.007
Teacher asked questions	0.136	-0.0269	-0.0634	0.213	-0.0205	-0.00630	394	0.014
Teacher asked apply info	0.103	-0.0306	-0.00648	0.0424	-0.0571**	-0.126**	394	0.016
Teacher tested creativity	0.271**	0.144**	0.0637	-0.0104	0.0925***	0.0667*	394	0.026
Teacher gave positive feedback	0.131*	0.103	0.160***	-0.0314	-0.0503*	-0.0291	394	0.022
Teacher gave corrective feedback	0.00507	0.0321	0.0759**	-0.0755	0.0540***	-0.0302	394	0.009
Teacher scolded pupils	0.0863	-0.0410***	0.0420	0.0112	-0.0700***	-0.155**	394	0.019
Teacher used local language	-0.0298	-0.00737	0.00781	-0.0728	-0.0563***	-0.0780*	394	0.022

Note: Levels of significance: *** p<0.01, ** p<0.05, * p<0.1.

Figure C 3. East African primary teachers' comparative mathematics performance



IX. REFERENCES

Dickerson, A., McIntosh, S., & Valente, C. (2015). Do the maths: An analysis of the gender gap in mathematics in Africa. *Economics of Education Review*, Vol. 46, pp. 1–22.

Government of Tanzania (2013). Education NKRA Lab Report: Executive summary. *Big Results Now*.

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review,* vol. 30: pp. 466–479.

Herrmann, M. A., & Rockoff, J. E. (2012). Worker absence and productivity: Evidence from teaching. *Journal of Labor Economics*, vol. 30(4), pp. 749-782.

Holmlund, H. & Sund, K. (2008). Is the gender gap in school performance affected by the sex of the teacher? *Labour Economics,* vol. 15: pp. 37–53.

Kitta, S. & Fussy, D. (2013). Bottlenecks in preparation of quality teachers in Tanzania. *Time Journals of Arts and Educational Research,* vol. 1(5) pp. 29-38.

Metzler, J. & Woessmann, L. (2012). The impact of teacher subject knowledge on pupil achievement: Evidence from within-teacher within-pupil variation. *Journal of Development Economics,* vol. 99: pp. 486–496.

Mott, M. S., Robinson, D. H., Walden, A., Burnette, J., & Rutherford, A.S. (2012). Illuminating the Effects of Dynamic Lighting on Pupil Learning. *SAGE Open*, pp. 1 –9.

PMO-RALG (2014). "Pre- Primary, Primary and Secondary Education Statistics in Brief 2013."

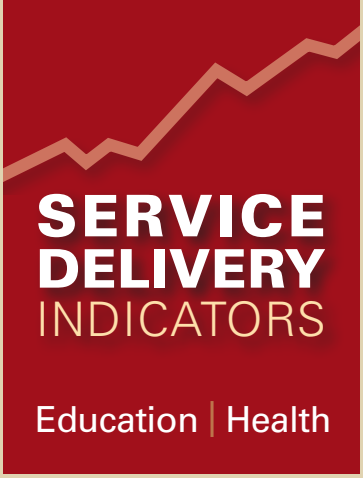
Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, vol. 73(2), pp.417–458.

Sabarwal, S., Evans, D., & Marshak, A. (2014). "The permanent input hypothesis: The case of textbooks and (No) pupil learning in Sierra Leone." *World Bank Policy Research Working Paper #7021*.

Tanzania Institute of Education (Ministry of Education and Vocational Training) (2007). "Curriculum for diploma in teacher education programmes in Tanzania."

World Bank (2014). "Tanzania Economic Update: Who Wants a Job? The Magnetic Power of Cities." *Report # 88932*.

World Bank. (2004). *World Development Report 2004: Making Services Work for Poor People*. World Bank. Washington, DC.



sdi@worldbank.org
www.worldbank.org/**SDI**
www.**SDI**ndicators.org

With support from The William and Flora Hewlett Foundation